

ESCUELA SUPERIOR POLITÉCNICA DEL LITORAL



FACULTAD DE CIENCIAS NATURALES Y MATEMÁTICAS

DEPARTAMENTO DE MATEMATICAS

PROYECTO DE TITULACIÓN

PREVIO A LA OBTENCIÓN DEL TÍTULO DE:

“MAGÍSTER EN ESTADÍSTICA APLICADA”

TEMA:

IMPLEMENTACIÓN DE ALGORITMOS CON APRENDIZAJE
REFORZADO PARA LA PLANIFICACIÓN DE LA CADENA DE
SUMINISTRO DE UNA EMPRESA DE BALANCEADO EN LA
CIUDAD DE GUAYAQUIL

AUTOR:

DANIEL ORLANDO CASTRO BURGOS

Guayaquil - Ecuador

2022

Resumen

Los algoritmos de aprendizaje estadístico tienen un ámbito de aplicación amplio, en el presente trabajo se aplicaron para definir la secuenciación de las órdenes de producción de una empresa de la industria de producción de alimentos para mascotas.

Inicialmente se implementó un algoritmo determinístico-estocástico generador de secuencias de producción a partir de las órdenes de producción reales, también se crearon variables adicionales asociadas a estas secuencias y luego se clasificaron las mismas en dos niveles; secuencia óptima y secuencia no óptima. Con estas variables se crearon dos conjuntos de datos: conjunto 1 contiene la secuencia de producción modelada como variables indicadoras y conjunto 2 sin estas.

A partir de los conjuntos creados se implementaron y ajustaron los algoritmos de clasificación: Máquina Vector Soporte y Árbol de Clasificación, finalmente con los modelos obtenidos se evaluó su desempeño en la partición de entrenamiento y prueba mediante las métricas de la matriz de confusión.

Finalmente se constató el rendimiento del algoritmo Máquina Vector Soporte es superior al del Árbol Clasificación para el conjunto 2. Para el conjunto 1, el desempeño del algoritmo Máquina Vector Soporte varía sustancialmente según el kernel. El desempeño del Algoritmo Árbol de Clasificación se mantiene estable en ambos conjuntos.

Palabras Claves: Secuenciación, Máquina vector soporte, Árbol Clasificación, Kernel, Métricas.

ABSTRACT

Statistical learning algorithms have a wide scope of application, in the present work they were applied to define the sequencing of production orders of a company in the pet food production industry.

Initially, a deterministic-stochastic algorithm was implemented to generate production sequences from the actual production orders, additional variables associated with these sequences were also created and then classified into two levels: optimal sequence and non-optimal sequence. With these variables, two data sets were created: set 1 contains the modeled production sequence as indicator variables and set 2 without these variables.

From the created sets, the classification algorithms were implemented and adjusted: Support Vector Machine and Classification Tree. Finally, with the obtained models, their performance in the training and test partition was evaluated by means of the confusion matrix metrics.

Finally, the performance of the Support Vector Machine algorithm was found to be superior to that of the Classification Tree for set 2. For set 1, the performance of the Support Vector Machine algorithm varies substantially depending on the kernel. The performance of the Classification Tree Algorithm remains stable for both sets.

Keywords: Sequencing, Support Vector Machine, Classification Tree, Kernel, Metrics.

DEDICATORIA

A DIOS y a mi familia.

AGRADECIMIENTO

Agradezco en primer lugar a mi familia por su incansable apoyo y comprensión en el tiempo de realización del presente trabajo.

También agradezco a mi tutor, Doctor Johny Pambabay por su guía y experiencia compartida.

DECLARACIÓN EXPRESA

La responsabilidad por los hechos y doctrinas expuestas en este Proyecto de Titulación me corresponde exclusivamente y ha sido desarrollado respetando derechos intelectuales de terceros conforme las citas que constan en el documento, cuyas fuentes se incorporan en las referencias o bibliografías. Consecuentemente este trabajo es de mi total autoría. El patrimonio intelectual del mismo corresponde exclusivamente a la ESCUELA SUPERIOR POLITÉCNICA DEL LITORAL.

En virtud de esta declaración, me responsabilizo del contenido, veracidad y alcance del Trabajo de Titulación referido.

Daniel Orlando Castro Burgos

TRIBUNAL DE GRADUACIÓN

Omar Honorio Ruiz Barzola Ph.D.

PRESIDENTE

Johny Javier Pambabay Calero Ph.D.

TUTOR

Francisco Antonio Moreira Villegas M. Sc.

DOCENTE EVALUADOR

TABLA DE CONTENIDO

CAPÍTULO 1.....	1
1. Introducción	1
1.1. Antecedentes.....	1
1.2. Descripción del problema	2
1.3. Objetivos.....	3
1.4. Hipótesis.....	4
1.5. Alcance.....	4
CAPÍTULO 2.....	5
2. Marco Teórico.....	5
2.1. Etapa 1: Conceptualización.....	5
2.2. Etapa 2: Revisión Bibliográfica.....	7
2.3. Etapa 3: Definición de Algoritmos.....	12
CAPÍTULO 3.....	21
3. Metodología	21
3.1. Fuente de Datos	21
3.2. Software	21
3.3. Metodología de Trabajo.....	22
3.4. Etapa de Importar.....	22

3.5.	Etapa Ordenar	23
3.6.	Etapa Transformar	24
3.7.	Etapa Visualización.....	27
3.8.	Etapa de Modelización.....	28
3.8.1.	Estimación de Funciones de Distribución de Probabilidad.....	28
3.8.2.	Algoritmo Generador de Secuencias	29
3.8.2.1.	Conjunto 1: Variables Predictoras: Indicadoras, Categóricas y Cuantitativas	33
3.8.2.2.	Conjunto 2: Variables Predictoras: Categóricas y Cuantitativas. 34	34
3.8.3.	Algoritmos de Clasificación	34
3.8.3.1.	Máquinas Vector Soporte	35
3.8.3.2.	Árbol de Clasificación	36
3.8.3.3.	Comparación de Algoritmos de clasificación.....	38
	CAPÍTULO 4	43
4.	Resultados	43
4.1.	Análisis Exploratorio Descriptivo.....	43
4.1.1.	Análisis Univariante Cuantitativo.....	43
4.1.2.	Análisis Univariante Categórico.	46
4.1.3.	Análisis Bivariante: Univariante Cuantitativo vs Multivariante Categórico 47	47

4.1.4. Análisis Bivariante: Bivariante Cuantitativo vs Multivariante Categórico	47
4.2. Estimación de Funciones de Distribución de Probabilidad.	49
4.3. Algoritmo Generador de Secuencias.	54
4.3.1. Criterio Mínimo Tiempo de Para entre el SKU's	54
4.3.2. Criterio Máxima Rapidez de SKU	55
4.4. Implementación de Algoritmos de Clasificación.	57
4.4.1. Máquina Vector Soporte	57
4.4.1.1. Conjunto 1: Estimación de Hiperparámetros C y γ.	57
4.4.1.2. Conjunto 1: Modelo	59
4.4.1.3. Conjunto 1: Matriz de Confusión.	59
4.4.1.4. Conjunto 2: Estimación de Hiperparámetros C y γ.	62
4.4.1.5. Conjunto 2: Modelo	64
4.4.1.6. Conjunto 2: Matriz de Confusión.	65
4.4.2. Árbol de Clasificación	66
4.4.2.1. Conjunto 1: Modelo	66
4.4.2.2. Conjunto 1: Matriz de Confusión.	69
4.4.2.3. Conjunto 2: Modelo	70
4.4.2.4. Conjunto 2: Matriz de Confusión.	70
4.4.3. Comparación de Algoritmos.	72
4.4.3.1. Métricas	72

4.4.3.2. Tiempos de ejecución	74
4.5. Evaluación de Algoritmos de Clasificación.	77
4.5.1. Máquina Vector Soporte	77
4.5.1.1. Conjunto 1: Matriz de Confusión.	77
4.5.1.2. Conjunto 2: Matriz de Confusión.	79
4.5.2. Árbol de Clasificación.....	80
4.5.2.1. Conjunto 1: Matriz de Confusión:	80
4.5.2.2. Conjunto 2: Matriz de Confusión.	81
4.5.3. Comparación de Algoritmos.....	83
CAPÍTULO 5.....	86
5. Conclusiones y Recomendaciones	86
5.1. Fortalezas:	86
5.2. Oportunidades	87
5.3. Debilidades	88
5.4. Amenazas	88
6. Referencias	89
Referencias	89
7. Apéndice y Anexos	91
7.1. Apéndice A: Tabla de Medidas de Tendencia Central Análisis Univariante Cuantitativo.....	91

7.2.	Apéndice B: Gráfico de barras de Análisis Univariante Categórico	
		100
7.3.	Apéndice B: Diagrama de Cajas de Análisis Bivariante: Univariante Cuantitativo Univariante vs Multivariante Categórico.	104
7.4.	Apéndice D: Gráfico de Dispersión de Análisis Bivariante: Bivariante Cuantitativo Univariante vs Multivariante Categórico.	117
7.5.	Apéndice E: Matrices de Confusión Máquina Vector Soporte Kernel Radial Ilustración 7: 67 Ilustración 7: 68 Ilustración 7: 69 Ilustración 7: 70	
		136
7.6.	Apéndice F: Matrices de Confusión Máquina Vector Soporte Kernel Polinomial Ilustración 7: 71 Ilustración 7: 72 Ilustración 7: 73 Ilustración 7: 74	
		138
7.7.	Apéndice G: Matrices de Confusión Árbol de Clasificación Ilustración 7: 75 Ilustración 7: 76 Ilustración 7: 77 Ilustración 7: 78	140

LISTADO DE ILUSTRACIONES

Ilustración 2. 1: Algoritmo KNN con $k=3$	15
Ilustración 2. 2: Algoritmo Árbol de Decisión.....	19
Ilustración 3. 1: Etapas de Proyecto.....	22
Ilustración 4: 1: Histograma Tiempo de Producción.....	44
Ilustración 4: 2: Histograma de Kilogramos.....	44
Ilustración 4: 3: Histograma Rapidez.....	45
Ilustración 4: 4: Histograma Tiempo de preparación.....	45
Ilustración 4: 5: Distribución Inversa Gaussiana Tiempo de Para.....	50
Ilustración 4: 6: Distribución Weibull Tiempo de Para.....	50
Ilustración 4: 7: Distribución Exponencial Tiempo de Para.....	51
Ilustración 4: 8: Frecuencia de Distribuciones de Tiempo de Para.....	51
Ilustración 4: 9: Distribución Weibull Rapidez.....	52
Ilustración 4: 10: Distribución Gamma Inversa Rapidez.....	52
Ilustración 4: 11: Distribución Gamma Rapidez.....	53

Ilustración 4: 12: Frecuencia de Distribuciones de Rapidez.....	53
Ilustración 4: 14: Error de Clasificación vs Hiperparámetros C y γ Kernel Radial Conjunto 1.	57
Ilustración 4: 15: Error de Clasificación vs Hiperparámetros C y γ Kernel Polinomial Conjunto 1.....	58
Ilustración 4: 16: Error de Clasificación vs Hiperparámetros C y γ Kernel Radial Conjunto 2.	62
Ilustración 4: 17: Error de Clasificación vs Hiperparámetros C y γ Kernel Polinomial Conjunto 2.....	63
Ilustración 4: 18: Parámetro c_p vs x_{error}	67
Ilustración 4: 19: Árbol de Clasificación.	68
Ilustración 4: 20: Exactitud Partición Entrenamiento por Algoritmo y Conjunto.	72
Ilustración 4: 21: Precisión Partición Entrenamiento por Algoritmo y Conjunto.	73
Ilustración 4: 22: Kappa Partición Entrenamiento por Algoritmo y Conjunto.	73
Ilustración 4: 23: Tiempo de ejecución Partición Entrenamiento por Algoritmo y Conjunto.....	75
Ilustración 4: 24: Exactitud Partición Prueba por Algoritmo y Conjunto.	83
Ilustración 4: 25: Precisión Partición Prueba por Algoritmo y Conjunto.	84

Ilustración 4: 26: Kappa Partición Prueba por Algoritmo y Conjunto.	84
Ilustración 7: 1: Gráfico de barras Presentación.....	101
Ilustración 7: 2: Gráfico de barras Tipo de Proceso.	101
Ilustración 7: 3: Gráfico de barras Turno.	102
Ilustración 7: 4: Gráfico de barras Día.....	102
Ilustración 7: 5: Gráfico de barras Mes.....	103
Ilustración 7: 6: Gráfico de barras Año.....	103
Ilustración 7: 7: Diagrama de Cajas Tiempo de Producción vs Presentación.	104
Ilustración 7: 8: Diagrama de Cajas Kilogramos vs Presentación.	105
Ilustración 7: 9: Diagrama de Cajas Rapidez vs Presentación.	105
Ilustración 7: 10: Diagrama de Cajas Tiempo Preparación vs Presentación.	106
Ilustración 7: 11: Diagrama de Cajas Tiempo de Producción vs Tipo de Proceso.....	106
Ilustración 7: 12: Diagrama de cajas Kilogramos vs Tipo de Proceso.....	107
Ilustración 7: 13: Diagrama de cajas Rapidez vs Tipo de Proceso.....	107
Ilustración 7: 14: Diagrama de cajas Tiempo de Preparación vs Tipo de Proceso.....	108

Ilustración 7: 15: Diagrama de cajas Tiempo Producción vs Turno.	108
Ilustración 7: 16: Diagrama de cajas Kilogramos vs Turno.	109
Ilustración 7: 17: Diagrama de cajas Rapidez vs Turno.	109
Ilustración 7: 18: Diagrama de cajas Tiempo de Preparación vs Turno. ..	110
Ilustración 7: 19: Diagrama de cajas Tiempo de Producción vs Día.	110
Ilustración 7: 20: Diagrama de cajas Kilogramos vs Día.	111
Ilustración 7: 21: Diagrama de cajas Rapidez vs Día.	111
Ilustración 7: 22: Diagrama de cajas Tiempo Preparación vs Día.	112
Ilustración 7: 23: Diagrama de cajas Tiempo Producción vs Mes.	112
Ilustración 7: 24: Diagrama de cajas Kilogramos vs Mes.	113
Ilustración 7: 25: Diagrama de cajas Rapidez vs Mes.	113
Ilustración 7: 26: Diagrama de cajas Tiempo Preparación vs Mes.	114
Ilustración 7: 27: Diagrama de cajas Tiempo Producción vs Año.	114
Ilustración 7: 28: Diagrama de cajas Kilogramos vs Año.	115
Ilustración 7: 29: Diagrama de cajas Rapidez vs Año.	115
Ilustración 7: 30: Diagrama de cajas Tiempo Preparación vs Año.	116
Ilustración 7: 31: Gráfico de Dispersión Tiempo de Producción vs Kilogramo vs Presentación.	117

Ilustración 7: 32: Gráfico de Dispersión Tiempo de Producción vs Rapidez vs Presentación.	118
Ilustración 7: 33: Gráfico de Dispersión Tiempo de Producción vs Tiempo de Preparación vs Presentación.	118
Ilustración 7: 34: Gráfico de Dispersión Kilogramos vs Rapidez vs Presentación.	119
Ilustración 7: 35: Gráfico de Dispersión Kilogramos vs Tiempo de Preparación vs Presentación.	119
Ilustración 7: 36: Gráfico de Dispersión Rapidez vs Tiempo de Preparación vs Presentación.....	120
Ilustración 7: 37: Gráfico de Dispersión Tiempo de Producción vs Kilogramo vs Tipo de Proceso.	120
Ilustración 7: 38: Gráfico de Dispersión Tiempo de Producción vs Rapidez vs Tipo de Proceso.	121
Ilustración 7: 39: Gráfico de Dispersión Tiempo de Producción vs Tiempo de Preparación vs Tipo de Proceso.	121
Ilustración 7: 40: Gráfico de Dispersión Kilogramos vs Rapidez vs Tipo de Proceso.....	122
Ilustración 7: 41: Gráfico de Dispersión Kilogramos vs Tiempo de Preparación vs Tipo de Proceso.	122

Ilustración 7: 42: Gráfico de Dispersión Rapidez vs Tiempo de Preparación vs Tipo de Proceso.	123
Ilustración 7: 43: Gráfico de Dispersión Tiempo de Producción vs Kilogramos vs Turno.....	123
Ilustración 7: 44: Gráfico de Dispersión Tiempo de Producción vs Rapidez vs Turno.	124
Ilustración 7: 45: Gráfico de Dispersión Tiempo de Producción vs Tiempo de Preparación vs Turno.....	124
Ilustración 7: 46: Gráfico de Dispersión Kilogramos vs Rapidez vs Turno.	125
Ilustración 7: 47: Gráfico de Dispersión Kilogramos vs Tiempo de Preparación vs Turno.....	125
Ilustración 7: 48: Gráfico de Dispersión Rapidez vs Tiempo de Preparación vs Turno.....	126
Ilustración 7: 49: Gráfico de Dispersión Tiempo de Producción vs Kilogramos vs Día.....	126
Ilustración 7: 50: Gráfico de Dispersión Tiempo de Producción vs Rapidez vs Día.	127
Ilustración 7: 51: Gráfico de Dispersión Tiempo de Producción vs Tiempo de Preparación vs Día.	127
Ilustración 7: 52: Gráfico de Dispersión Kilogramos vs Rapidez vs Día... ..	128

Ilustración 7: 53: Gráfico de Dispersión Kilogramos vs Tiempo de Preparación vs Día.....	128
Ilustración 7: 54: Gráfico de Dispersión Rapidez vs Tiempo de Preparación vs Día.....	129
Ilustración 7: 55: Gráfico de Dispersión Tiempo de Producción vs Kilogramo vs Mes.....	129
Ilustración 7: 56: Gráfico de Dispersión Tiempo de Producción vs Rapidez vs Mes.	130
Ilustración 7: 57: Gráfico de Dispersión Tiempo de Producción vs Tiempo de Preparación vs Mes.	130
Ilustración 7: 58: Gráfico de Dispersión Kilogramos vs Rapidez vs Mes. .	131
Ilustración 7: 59: Gráfico de Dispersión Kilogramos vs Tiempo de Preparación vs Mes.	131
Ilustración 7: 60: Gráfico de Dispersión Rapidez vs Tiempo de Preparación vs Mes.....	132
Ilustración 7: 61: Gráfico de Dispersión Tiempo de Producción vs Kilogramo vs Año.	132
Ilustración 7: 62: Gráfico de Dispersión Tiempo de Producción vs Rapidez vs Año.....	133
Ilustración 7: 63: Gráfico de Dispersión Tiempo de Producción vs Tiempo de Preparación vs Año.....	133

Ilustración 7: 64: Gráfico de Dispersión Kilogramo vs Rapidez vs Año....	134
Ilustración 7: 65: Gráfico de Dispersión Kilogramo vs Tiempo de Preparación vs Año.....	134
<i>Ilustración 7: 66: Gráfico de Dispersión Rapidez vs Tiempo de Preparación vs Año.....</i>	<i>135</i>
Ilustración 7: 67: Matriz de Confusión Kernel Radial Partición Entrenamiento Conjunto 1	136
Ilustración 7: 68: Matriz de Confusión Kernel Radial Partición Entrenamiento Conjunto 2	136
Ilustración 7: 69: Matriz de Confusión Kernel Radial Partición Prueba Conjunto 1	137
Ilustración 7: 70: Matriz de Confusión Kernel Radial Partición Prueba Conjunto 2	137
Ilustración 7: 71: Matriz de Confusión Kernel Polinomial Partición Entrenamiento Conjunto 1	138
Ilustración 7: 72: Matriz de Confusión Kernel Polinomial Partición Entrenamiento Conjunto 2	138
Ilustración 7: 73: Matriz de Confusión Kernel Polinomial Partición Prueba Conjunto 1	139
Ilustración 7: 74: Matriz de Confusión Kernel Polinomial Partición Prueba Conjunto 2	139

Ilustración 7: 75: Matriz de Confusión Árbol de Clasificación Partición Entrenamiento Conjunto 1	140
Ilustración 7: 76: Matriz de Confusión Árbol de Clasificación Partición Entrenamiento Conjunto 2	140
Ilustración 7: 77: Matriz de Confusión Árbol de Clasificación Partición Prueba Conjunto 1.....	141
Ilustración 7: 78: Matriz de Confusión Árbol de Clasificación Partición Prueba Conjunto 2.....	141

Listado de Tablas

Tabla 3. 1: Variables iniciales	23
Tabla 3. 2: Variables excluidas.	24
Tabla 3. 3: Variables creadas.	25
Tabla 3. 4: Variables finales.	26
Tabla 3. 5: Matriz de Estados.....	30
Tabla 3. 6: Matriz de Confusión	39
Tabla 3. 7: Métricas.....	40
Tabla 3. 8: Interpretación Estadístico Kappa	41
Tabla 3. 9: Estadísticos	41
Tabla 4. 1: Secuencias Criterio Mínimo Tiempo de Para.	55
Tabla 4. 2: Secuencias Criterio Máxima Rapidez de Producción.....	56
Tabla 4. 3: Comparación Secuencias Min Tiempo Para vs Max Rapidez..	56
Tabla 4. 4: Hiperparámetros Máquina Vector Soporte Kernel Radial y Polinomial Conjunto 1.....	58
Tabla 4. 5: Vectores Soporte Kernel Radial y Polinomial Conjunto 1.....	59
Tabla 4. 6: Matriz de Confusión Partición Entrenamiento Kernel Radial Conjunto 1.	59

Tabla 4. 7: Matriz de Confusión Partición Entrenamiento Kernel Polinomial Conjunto 1.....	60
Tabla 4. 8: Métricas Partición Entrenamiento Máquina Vector Soporte Conjunto 1.....	60
Tabla 4. 9: Estadísticos Partición Entrenamiento Conjunto 1.....	61
Tabla 4. 10: Hiperparámetros Máquina Vector Soporte Kernel Radial y Polinomial Conjunto 2.	63
Tabla 4. 11: Vectores Soporte Kernel Radial y Polinomial Conjunto 2.	64
Tabla 4. 12: Matriz de Confusión Partición Entrenamiento Kernel Radial Conjunto 2.....	65
Tabla 4. 13: Matriz de Confusión Partición Entrenamiento Kernel Polinomial Conjunto 2.....	65
Tabla 4. 14: Métricas Partición Entrenamiento Máquina Vector Soporte Conjunto 2.....	65
Tabla 4. 15: Estadísticos Partición Entrenamiento Conjunto 2.....	66
Tabla 4. 16: Importancia de Variables Conjunto 1.....	66
Tabla 4. 17: Reglas de Árbol de Clasificación.....	68
Tabla 4. 18: Matriz de Confusión Partición Entrenamiento Árbol de Clasificación Conjunto 1.....	69
Tabla 4. 19: Métricas Partición Entrenamiento Conjunto 1.....	69

Tabla 4. 20: Estadísticos Partición Entrenamiento Conjunto 1.	70
Tabla 4. 21: Importancia de Variables Conjunto 2.	70
Tabla 4. 22: Matriz de Confusión Partición Entrenamiento Árbol de Clasificación Conjunto 2.	70
Tabla 4. 23: Métricas Partición Entrenamiento Conjunto 2	71
Tabla 4. 24: Estadísticos Partición Entrenamiento Conjunto 2.	71
Tabla 4. 25: Métricas Partición Entrenamiento por Algoritmo y Conjunto ..	72
Tabla 4. 26: Tiempo de ejecución Partición Entrenamiento por Algoritmo y Conjunto.	75
Tabla 4. 27: Matriz de Confusión Partición Prueba Kernel Radial Conjunto 1.	77
Tabla 4. 28: Matriz de Confusión Partición Prueba Kernel Polinomial Conjunto 1.	77
Tabla 4. 29: Métricas Partición Prueba Máquina Vector Soporte Conjunto 1.	78
Tabla 4. 30: Estadísticos Partición Prueba Máquina Vector Soporte Conjunto 1.....	78
Tabla 4. 31: Matriz de Confusión Partición Prueba Kernel Radial Conjunto 2.	79
Tabla 4. 32: Matriz de Confusión Partición Prueba Kernel Polinomial Conjunto 2.	79

Tabla 4. 33: Métricas Partición Prueba Máquina Vector Soporte Conjunto 2.	79
Tabla 4. 34: Estadísticos Partición Prueba Máquina Vector Soporte Conjunto 1.....	80
Tabla 4. 35: Matriz de Confusión Partición Prueba Árbol de Clasificación Conjunto 1.....	80
Tabla 4. 36: Matriz de Confusión Partición Prueba Árbol de Clasificación Conjunto 1.....	81
Tabla 4. 37: Estadísticos Partición Prueba Árbol de Clasificación Conjunto 1.	81
Tabla 4. 38: Matriz de Confusión Partición Prueba Árbol de Clasificación Conjunto 2.....	81
Tabla 4. 39: Métricas Partición Prueba Árbol de Clasificación Conjunto 2.	82
Tabla 4. 40: Estadísticos Partición Prueba Árbol de Clasificación Conjunto 2.	82
Tabla 4. 41: Métricas Partición Prueba por Algoritmo y Conjunto.	83
Tabla 7. 1: Medidas de Tendencia Central y Medias de Dispersión: Presentación Ensacado.	91
Tabla 7. 2: Medidas de Tendencia Central y Medias de Dispersión: Presentación Granel.	92

Tabla 7. 3: Medidas de Tendencia Central y Medias de Dispersión: Presentación Reproceso.....	92
Tabla 7. 4: Medidas de Tendencia Central y Medias de Dispersión: Tipo Proceso Polvo.....	92
Tabla 7. 5: Medidas de Tendencia Central y Medias de Dispersión: Tipo Proceso Pellet.....	93
Tabla 7. 6: Medidas de Tendencia Central y Medias de Dispersión: Tipo Proceso Expandido.....	93
Tabla 7. 7: Medidas de Tendencia Central y Medias de Dispersión: Turno Diurno.....	93
Tabla 7. 8: Medidas de Tendencia Central y Medias de Dispersión: Turno Nocturno.....	94
Tabla 7. 9; Medidas de Tendencia Central y Medias de Dispersión: lunes.	94
Tabla 7. 10: Medidas de Tendencia Central y Medias de Dispersión: martes	94
Tabla 7. 11: Medidas de Tendencia Central y Medias de Dispersión: miércoles	95
Tabla 7. 12: Medidas de Tendencia Central y Medias de Dispersión: jueves	95

Tabla 7. 13: Medidas de Tendencia Central y Medias de Dispersión: viernes	95
Tabla 7. 14: Medidas de Tendencia Central y Medias de Dispersión: sábado	95
Tabla 7. 15: Medidas de Tendencia Central y Medias de Dispersión: domingo	96
Tabla 7. 16: Medidas de Tendencia Central y Medias de Dispersión: Mes Enero.	96
Tabla 7. 17: Medidas de Tendencia Central y Medias de Dispersión: Mes Febrero.	96
Tabla 7. 18: Medidas de Tendencia Central y Medias de Dispersión: Mes Marzo.	96
Tabla 7. 19: Medidas de Tendencia Central y Medias de Dispersión: Mes Abril.	97
Tabla 7. 20: Medidas de Tendencia Central y Medias de Dispersión: Mes Mayo.	97
Tabla 7. 21: Medidas de Tendencia Central y Medias de Dispersión: Mes Junio.	97
Tabla 7. 22: Medidas de Tendencia Central y Medias de Dispersión: Mes Julio.	97

Tabla 7. 23: Medidas de Tendencia Central y Medias de Dispersión: Mes Agosto.....	98
Tabla 7. 24: Medidas de Tendencia Central y Medias de Dispersión: Mes Septiembre.	98
Tabla 7. 25: Medidas de Tendencia Central y Medias de Dispersión: Mes Octubre.....	98
Tabla 7. 26: Medidas de Tendencia Central y Medias de Dispersión: Mes Noviembre.	98
Tabla 7. 27: Medidas de Tendencia Central y Medias de Dispersión: Mes Diciembre.....	99
Tabla 7. 28: Medidas de Tendencia Central y Medias de Dispersión: Año 2015.....	99
Tabla 7. 29: Medidas de Tendencia Central y Medias de Dispersión: Año 2016.....	99
Tabla 7. 30: Medidas de Tendencia Central y Medias de Dispersión: Año 2017.....	99
Tabla 7. 31: Medidas de Tendencia Central y Medias de Dispersión: Año 2018.....	100

Listado de Fórmulas

Fórmula 2. 1: Fórmula de Distancias Algoritmo KNN.	12
Fórmula 2. 2: Distancia de dos observaciones.	13
Fórmula 2. 3: Función partición y corte.	15
Fórmula 2. 4: Estimación de y	16
Fórmula 2. 5: Partición de la región en semi planos.	16
Fórmula 2. 6: Elección de variable a particionar.	16
Fórmula 2. 7: Estimación de c_1	16
Fórmula 2. 8: Cardinalidad de la región.	17
Fórmula 2. 9: Definición de cm	18
Fórmula 2. 10: Criterio Costo-Complejidad.	18
Fórmula 2. 11: Algoritmo Q-Learning	19
Fórmula 3. 1: Actualización tiempo acumulado con SKU inicial.	31
Fórmula 3. 2: Actualización tiempo acumulado con tiempo de para.	32
Fórmula 3. 3: Elección Secuencia Óptima.	32
Fórmula 3. 4: Variable etiqueta.	33
Fórmula 3. 5: Variables indicadoras.	33

Fórmula 3. 6: Producto Interno Vectorial.....	35
Fórmula 3. 7: Kernel Polinomial	35
Fórmula 3. 8: Kernel Radial	36
Fórmula 3. 9: Métrica Tasa de Error de Clasificación.	37
Fórmula 3. 10: Métrica Índice Gini	37
Fórmula 3. 11: Métrica Entropía Cruzada	37
Fórmula 3. 12: Deviance	39
Fórmula 3. 13: Estadístico Kappa.	41

CAPÍTULO 1

1. Introducción

1.1. Antecedentes

Para el año 2020 la Encuesta Global sobre Alimento Balanceado (ALLTECH, 2020) reveló que la estimación de la producción mundial de alimento balanceados fue de 1126 millones de toneladas métricas, de este total, países como Estados Unidos, China, Brasil, India, México, España, Japón y Alemania representan el 58% de la producción mundial. El 2% de esta producción es alimento balanceado para mascotas, siendo este uno de los tipos de balanceado de mayor crecimiento.

A nivel regional América Latina se ubica en la segunda posición entre las regiones con mayor crecimiento de la producción de balanceado para mascotas con un 6% de crecimiento superior al año 2019, solo por detrás de la región de Asia Pacífico que reporta un crecimiento del 10%. En la región de América Latina, los países que más aumentado la producción con Ecuador y Argentina.

Según PRO-ECUADOR (PROECUADOR, 2020) para el 2021 en Ecuador existen 351 empresas productoras de balanceado, estas a su vez forman parte de dos grandes asociaciones; la Asociación Ecuatoriana de Fabricantes de Alimentos Balanceados para Animales (AFABA) y la Asociación de Productores de Alimentos Balanceados (APROBAL)

Al igual que la tendencia mundial, en el Ecuador la demanda de balanceado de mascotas ha representado para las empresas una línea de negocio muy rentable,

llegando a representar el 24% de las ventas del año 2020 en una de las empresas líderes en la producción de balanceado de animales (Revista Vistazo, 2020)

A diferencia de balanceados producidos para los sectores de aves, porcino, bovino, acuicultura, etc., la producción de balanceados para mascotas tiene incorporada biotecnología *extralife* que permite mejorar la calidad de vida de las mascotas mediante la alimentación; además incorpora consideraciones como tipo de raza, tamaño, edad, todos estos factores únicos que diferencian este tipo de balanceados de los demás.

1.2. Descripción del problema

Planificar la cadena de suministro es un proceso complejo que comprende todas las actividades relacionadas al flujo de transformación de bienes, desde su adquisición, almacenamiento, tratamiento productivo y distribución a los consumidores; a lo largo de todo este flujo intervienen muchas variables tanto determinísticas como capacidad instalada, así como variables estocásticas, por lo cual es de vital importancia la optimización de la planificación de la cadena de suministro.

El efecto del aumento del costo de producción se puede traducir en un incremento del precio de venta para mantener el margen de ganancia adecuado o reducción de ganancia debido a la alta elasticidad de precios; en el primer escenario se pierde competitividad por precio mientras que en el segundo se pierde rentabilidad.

Entre los principales ejes de reducción de costo se encuentra el establecimiento de políticas óptimas de aprovisionamiento que permitan minimizar

costos de mantenimiento y almacenamiento de inventario, aquí el agente participante es el planificador de compras, mismo que debe establecer un equilibrio que permita satisfacer las necesidades de producción, pero sin incurrir en altos costos asociados al inventario.

Otro eje es el establecimiento de planificaciones de producción inteligentes que permitan maximizar la producción haciendo uso eficiente de los recursos disponibles tal que cumplan los objetivos de cumplir con la demanda y minimizar el costo unitario de producción. El planificador de la producción debe establecer planificaciones que permitan cumplir con la demanda sin incurrir en altos costos que repercuten en el costo del producto.

Por lo antes indicado, es claro que existe un problema en la planificación de la cadena de suministro, ante lo cual se requieren modelos de optimización de aprendizaje reforzado que permitan minimizar el costo de producción.

1.3. Objetivos

Objetivo General

Evaluar la planificación de la cadena de suministro de una empresa productora de balanceado para animales mediante la ejecución de algoritmos híbridos de aprendizaje reforzado utilizando software estadístico especializado.

Objetivos Específicos

Aplicar un algoritmo mediante modelización estadística para la planificación de la producción, sujeta a la demanda de los clientes.

Planear la secuenciación de la producción mediante la implementación de aprendizaje reforzado para satisfacer la demanda establecida.

Elegir el algoritmo adecuado mediante modelización estadística para el establecimiento de una política secuenciación de la producción.

1.4. Hipótesis

La implementación de algoritmos de aprendizaje estadístico ayuda la optimización de la planificación de la producción de alimentos balanceados para mascotas.

1.5. Alcance

Se implementarán algoritmos híbridos matemáticos-estadísticos para planificar la programación de la producción en una empresa de consumo masivo, productora de balanceados para animales localizada en la ciudad de Guayaquil, considerando los datos de los años 2015, 2016, 2017 y 2018.

CAPÍTULO 2

2. Marco Teórico

El Marco Teórico como sustento de la investigación se lo abordará en 3 etapas, las que son:

Etap 1: Conceptualización general del aprendizaje estadístico, fundamentos y objetivos.

Etap 2: Revisión de bibliografía relevante en plataformas como SCOPUS y Web of Science y Google Scholar con el objetivo de detallar casos de aplicación de diferentes algoritmos de aprendizaje estadístico en la cadena de suministro.

Etap 3: Definición de los algoritmos de aprendizaje estadístico implementados en la fase de revisión bibliográfica.

2.1. Etapa 1: Conceptualización.

El Aprendizaje Estadístico (Statistical Learning-SL en inglés) son herramientas cuyo objetivo principal es obtener la mayor cantidad de conocimiento de un conjunto de datos. El Aprendizaje Estadístico surgió como una rama de la estadística y se enfoca en la modelización estocásticas y su interpretación. Así mismo estudia la medición de la eficacia mediante el equilibrio de Exactitud y Precisión (sesgo y varianza)

De manera general las técnicas de Aprendizaje Estadístico se pueden clasificar en tres grandes grupos: Supervisado, No Supervisado y Reforzado.

El Aprendizaje Supervisado se identifica por la existencia de una variable respuesta Y denominada dependiente, así mismo se necesita un conjunto de variables predictivas $X = \{x_1, x_2, \dots, x_n\}$ también llamadas dependientes. Esta variable Y puede ser continua o no serlo, en el primer caso se considera que es un problema de Regresión, por ejemplo, el sueldo, peso, etc. de no ser continua se dice que es un problema de Clasificación y esta a su vez puede ser binaria (se otorga o no el crédito a un cliente que lo ha solicitado), nominal (color de ojos)

El Aprendizaje No Supervisado no tiene una variable a predecir, el interés en estos algoritmos radica en poder realizar asociaciones entre los individuos de un conjunto de datos cuando no existe una categorización predefinida de los mismos.

Un ejemplo de Aprendizaje No Supervisado es el caso de un supermercado que desea segmentar sus clientes, pero al ser los mismos cuasi homogéneos en su comportamiento de compra no es posible identificar grupos en los cuales asignar a los clientes. Este es el rango de aplicación de algoritmos de aprendizaje no supervisado considerando como inputs variables como frecuencia de compra, valor de compra, unidades compradas, tipo de producto de compra el cliente, etc., dando el agrupamiento de clientes con similares comportamientos de compra.

El Aprendizaje Reforzado se basa en recompensas por comportamientos deseados y penalizaciones por comportamientos no deseados. Este tipo de aprendizaje a diferencia del aprendizaje supervisado no existe una respuesta que el algoritmo debe alcanzar, en este caso el refuerzo decide como responder a una tarea según las recompensas.

En el aprendizaje reforzado existe un conjunto de algoritmos que se adaptan a la resolución de diferentes problemas, es decir algoritmos como *Q – learning* se pueden implementar para aprender políticas de inventario y también para poder definir la ruta óptima entre dos puntos dada una serie de obstáculos. En el primer caso la penalidad sería la infracción de capacidad de almacenamiento y en el segundo caso sería los obstáculos de la ruta.

Como es de notar el rango de aplicación de los algoritmos de aprendizaje reforzado es mayor al de los algoritmos de aprendizaje supervisado y no supervisado y es por esto el creciente ámbito de aplicación de estos en la industria.

2.2. Etapa 2: Revisión Bibliográfica.

Ante la gran variedad de ámbitos de aplicación del aprendizaje estadístico la misma ha incursionado en los procesos de producción permitiendo la optimización de estos, proceso que a la vez redundando en la reducción de costos y tiempos.

En el sector producción la aplicación de Aprendizaje Estadístico se enfoca en encontrar patrones a partir del conjunto de datos existentes de la producción para generar aproximaciones sobre el comportamiento futuro del sistema, otorgando una ventaja competitiva a los dueños del proceso al tomar decisiones. (Wuest, Weimer, Irgens, & Klaus-Dieter, 2016)

(Lieber, Stolpe, Konrad, Deuse, & Morik, 2013) toma la producción siderúrgica que se caracteriza por sistemas de producción altamente interconectados y complejos que absorben muchos recursos. Esta alta complejidad vuelve costoso el control de calidad de las características físicas de los productos, por este motivo las

desviaciones de calidad deben ser detectadas y corregidas a lo largo de la producción sin medidas intrusivas, por lo que se vuelve imprescindible mediante minería de datos de las mediciones de los sensores detectar las desviaciones de calidad lo antes posible y en tiempo real para evitar incurrir en reprocesos o rechazos de los productos terminados.

(Lieber, Stolpe, Konrad, Deuse, & Morik, 2013) implementa el algoritmo KNN para predecir la calidad (aprobada o rechazada) de las piezas basándose en los datos recopilados por los sensores. Toma 218 características que se miden a lo largo de los 470 procesos por los que discurre el proceso de producción.

Estas características son sugeridas por los expertos del proceso, entre las relevantes para la calidad se incluyen tiempo de calentamiento en el horno, fuerza de laminación, velocidad, temperatura, etc.

Se logra una calibración del algoritmo con $K=7$, es decir se clasifica los individuos según la clasificación de sus 7 vecinos más cercanos, se realizaron 50 corridas de la producción llegando a un 80.21% de precisión de la predicción.

Se ha esbozado la aplicación de un algoritmo de aprendizaje estadístico de clasificación, pero también existen algoritmos de aprendizaje estadístico de regresión, a continuación, se trata la aplicación de uno de estos a la producción.

Ante una integración vertical cada vez mayor en todas las industrias se ha aumentado la colaboración entre clientes-proveedores llegando a evolucionar esta relación a socios de negocios ya que es justamente factores como compartir

información en tiempo real de los niveles de inventario de los clientes de una fábrica, permite por un lado al cliente minimizar el tiempo de registro y colocación de pedidos a la fábrica y por otro lado permite a la fábrica tener una mejor planificación de la producción y evitar incurrir en superávit o déficit de la producción.

(Reboiro-Jato, y otros, 2011) implementaron algoritmos híbridos para una fábrica productora de alimento de aves que ha logrado integrarse verticalmente con sus socios que son las granjas de aves, la criticidad de esta integración está en que las aves se alimentan de manera continua por lo que los niveles de inventario se reducen de manera constante y a alta velocidad. Un déficit de alimento provoca un aumento de la mortalidad de las aves mismo que conlleva a pérdidas.

La metodología del trabajo consideró un modelo híbrido constituido por tres subsistemas:

- I. Recopilación de datos e identificación del tipo de alimentación.
- II. Previsión del consumo.
- III. Planificación de la producción.

El subsistema 1 considera variables como edad de las aves, tipo de aves entre otras, pero la identificación del tipo de alimentación es determinística.

El subsistema 2 estima la cantidad de cada tipo de alimento que requieren los diferentes lotes de aves, para la predicción del volumen de alimento se ejecuta un árbol de decisión de regresión utilizando datos históricos de la base interna.

El subsistema 3 resume y combina la información de los subsistemas 1 y 2 y determina la cantidad de alimento que producirá por día. Este subsistema simula la evolución de los silos de cada granja aumentando el inventario cada vez que se realiza una entrega a una granja.

Hasta el momento se han revisado casos de aplicación de algoritmos de aprendizaje estadístico supervisado y no supervisado, ahora el enfoque será en algoritmos de aprendizaje reforzado, estos algoritmos se han aplicado en distintos ámbitos de la industria.

Se han implementado estos algoritmos para resolver problemas de los siguientes tipos:

- I. Problemas NP hard
- II. Problemas en que no es factible aprendizaje supervisado ni no supervisado.

Los problemas NP hard son aquellos para los cuales existe un modelo matemático de optimización definido y se puede resolver mediante el algoritmo simplex, pero debido a su alta complejidad no pueden ser resueltos en un tiempo razonable.

Si bien es cierto para problemas NP hard se han desarrollado metaheurísticas que permiten obtener soluciones buenas, la aplicación de algoritmos de aprendizaje reforzado irrumpe como una herramienta novedosa para resolver estos problemas desde un enfoque estocástico y no determinístico.

Para el segundo tipo de problema, estos algoritmos abordan problemas para los cuales los algoritmos supervisados y no supervisados son de compleja implementación, sus resultados no son satisfactorios o no es posible modelar.

En el ámbito de aplicación de algoritmos de aprendizaje estadístico se encuentra la aportación de (Ahn, Park, Park, & Hur, 2019) que aplican un algoritmo *Q – learning* para el troquelado de semiconductores.

El proceso de pick-and-place es un problema de optimización muy conocido y es uno de los denominados problemas NP-hard, debido a esta particularidad es que se opta aplicar un algoritmo *Q – learning*. Se desarrolla un algoritmo en que interactúan los procesos de selección y ubicación de piezas y estos a su vez se retroalimentan entre sí de forma interactiva.

Hasta el momento se han revisado aplicaciones de algoritmos de aprendizaje estadístico supervisado, no supervisado y reforzado y su desempeño a la hora de ayudar a resolver diferentes problemas de la industria de la producción, pero es importante denotar que estos algoritmos por sí solo no pueden llegar a modelar en un 100% los problemas analizados. Por este motivo es importante la implementación de algoritmos híbridos que permitan obtener la potencia de los algoritmos descritos previamente, pero sin descuidar la ventaja de métodos determinísticos a la hora de resolver partes del problema que no pueden ser modelizados matemáticamente.

El enfoque en proyecto de fin de máster es evidenciar las ventajas de este tipo de metodología a la hora de abordar problemas de la industria.

2.3. Etapa 3: Definición de Algoritmos.

Algoritmo K Nearest Neighbours (KNN) o Algoritmo de los K Vecinos más cercanos.

Es un algoritmo de aprendizaje estadístico supervisado de clasificación que parte de un conjunto de datos de entrenamiento, cada dato dentro de una clase Su ejecución sigue los siguientes pasos:

Paso 1: Para cada individuo se mide su distancia a los demás individuos.

Paso 2: Ordena las distancias de menor a mayor.

Paso 3: Asigna el individuo a la clase más frecuente en la que se encuentran sus k vecinos más cercanos del listado del Paso 2.

Para la implementación del algoritmo es importante definir la estructura necesaria del conjunto de datos:

Sea $D_{m,n}$ una matriz que representa el conjunto de datos, esta se ilustra en

Fórmula 2. 1

$$D_{m,n} = \begin{pmatrix} d_1 & = & a_{1,1} & \cdots & a_{1,n} & c_1 \\ d_2 & = & a_{2,1} & \ddots & a_{2,n} & c_2 \\ \vdots & \vdots & \vdots & \cdots & \vdots & \vdots \\ d_m & = & a_{m,1} & \cdots & a_{m,n} & c_m \end{pmatrix}$$

Fórmula 2. 1: Fórmula de Distancias Algoritmo KNN.

Donde d_m representa al emésimo individuo sobre los que se realizaron las observaciones, $a_{m,n}$ indica el enésimo atributo del emésimo individuo y c_m es la clase del emésimo individuo.

Para la medición de la distancia existen varias métricas, una métrica clásica usada en KNN es la Métrica Euclidiana que es la distancia que se mediría al unir dos puntos mediante un segmento de recta, es decir si se deseara calcular la distancia euclidiana entre el individuo 1 y el individuo 2 estaría dada por Fórmula 2. 2

$$dist(d_1, d_2) = \sqrt{(a_{1,1} - a_{2,1})^2 + (a_{1,2} - a_{2,2})^2 + \dots + (a_{1,n} - a_{2,n})^2}$$

Fórmula 2. 2: Distancia de dos observaciones.

Hasta el momento se ha mencionado que el paso final del algoritmo KNN es asignar el individuo a la clase a la que pertenece la mayoría de sus K vecinos más cercanos, pero no hemos definido el valor de K o el método de escoger un K apropiado.

Es evidente que si $K=1$, entonces el individuo será asignado a la clase de su único vecino más cercano, por lo que sería tentador definir un valor de $K=1$ ya que este aseguraría que cada individuo se asigne a la misma clase del individuo más cercano, pero en este caso el modelo estaría expuesto a datos ruidosos o datos atípicos. Si el investigador al momento de tomar la medición de la clase de un individuo lo clasificó erróneamente, de manera segura este error provocará que el modelo clasifique erróneamente al individuo, es decir se pierde la información que otorgaría el segundo, tercero, etc. vecino más cercano al ser $K=1$.

Por otro lado, en el otro extremo sabemos que K podría tomar a lo sumo el valor de $m - 1$, de esta manera se estaría evitando verse afectado por valores atípicos o datos ruidosos, pero al momento de clasificar a un individuo este

pertenecerá de manera obligatoria a la clase más común, lo mismo pasará al momento de querer clasificar al segundo individuo y así de manera general se terminará por clasificar a todos los individuos en la clase más común.

Ahora que se sabe que los 2 extremos no garantizan un K tal que maximice la efectividad del algoritmo, se puede intuir que el mismo se encuentra entre estos dos valores.

Generalmente el valor de k depende de la dificultad de la clasificación que debe realizar el modelo, pero se suele ubicar entre 3 y 10, también es una práctica definirlo como la raíz cuadrada del número de individuos que forman parte del conjunto de entrenamiento así mismo gracias a la potencia que otorgan los ordenadores se pueden crear múltiples conjuntos de entrenamiento y sobre estos probar varios valores de K y finalmente elegir el que obtenga mejor rendimiento en la clasificación (Hastie, Tibshirani, & Friedman, 2018)

En Ilustración 2. 1 se puede observar un ejemplo de la ejecución de un algoritmo KNN donde $k=3$, como se puede observar el punto verde toma sus 3 vecinos más cercanos los cuales son: 2 puntos azules y un 1 naranja. El algoritmo clasifica al punto verde como verde ya que es el más común entre sus 3 vecinos.

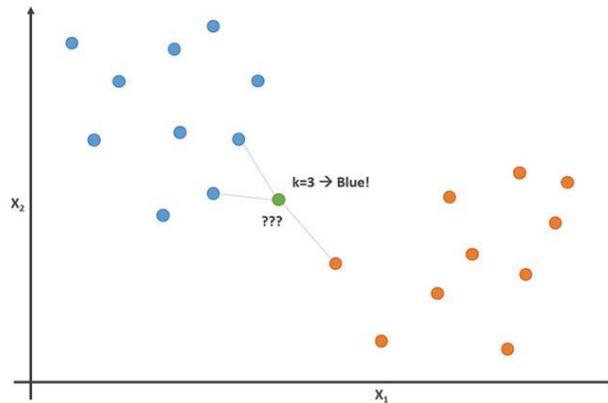


Ilustración 2. 1: Algoritmo KNN con $k=3$.

Algoritmo Árbol de Decisión.

Para (James, Witten, Hastie, & Tibshirami, 2013) es un algoritmo de aprendizaje estadístico supervisado que se denomina de regresión cuando la variable respuesta Y es continua o de clasificación cuando la variable respuesta y es categórica. Los fundamentos de un árbol de decisión son los siguientes:

Se considera un espacio de N características donde cada uno tiene su correspondiente variable respuesta $Y = \{Y_1, Y_2, \dots, Y_N\}$, sobre este conjunto de datos el algoritmo debe decidir automáticamente la variable de división, así como el punto de división y la topología que el árbol debe tener. Se asume que inicialmente se tiene una partición de M regiones denotadas por R_1, R_2, \dots, R_M y se modela la respuesta como una constante c_m en cada región mediante Fórmula 2. 3

$$f(x) = \sum_{m=1}^M c_m I(x \in R_m)$$

Fórmula 2. 3: Función partición y corte.

Donde es de interés estimar c_m ya que es el valor que tomará x cuando esta pertenece a la región R_m . Se puede evidenciar que una buena estimación de c_m sería el promedio de la variable respuesta de todos los individuos que forman parte de la región R_m , usando la Fórmula 2. 4

$$\widehat{c}_m = AVE(y_i | x_i \in R_m)$$

Fórmula 2. 4: Estimación de y.

Ahora, el particionamiento se lo realiza mediante un algoritmo voraz donde se elige una variable inicial de división j y un punto s de división de la región en dos semi planos, según la Fórmula 2. 5

$$R_1(j, s) = \{X | X_j \leq s\} \quad \wedge \quad R_2(j, s) = \{X | X_j > s\}$$

Fórmula 2. 5: Partición de la región en semi planos.

Luego se busca la variable de división j y el punto de división s que resuelven conforme la Fórmula 2. 6

$$\min_{j,s} \left[\min_{c_1} \sum_{x_i \in R_1(j,s)} (y_i - c_1)^2 + \min_{c_2} \sum_{x_i \in R_2(j,s)} (y_i - c_2)^2 \right]$$

Fórmula 2. 6: Elección de variable a particionar

Para cualquier elección j y s , la minimización interna se resuelve mediante la Fórmula 2. 7

$$\widehat{c}_1 = AVE(y_i | x_i \in R_1(j, s)) \quad \wedge \quad AVE(y_i | x_i \in R_2(j, s))$$

Fórmula 2. 7: Estimación de \widehat{c}_1

Una vez encontrada la mejor división, se procede a dividir los datos en dos regiones, luego se repite el proceso en todas las regiones resultantes y así sucesivamente.

Una importante pregunta para responder es ¿Cuánto debe crecer el árbol mediante las ramificaciones? Es evidente que si el árbol se expande demasiado se podría incurrir en un sobreajuste de los datos, pero si solo se generan pocas ramificaciones se puede tener un ajuste de poca calidad. El tamaño del árbol es un parámetro por estimar, se pueden usar dos criterios:

Criterio 1: Se realiza una partición solo si la diferencia entre la suma de cuadrados entre el nivel actual y el siguiente supera un umbral predefinido.

Criterio 2: Se establece arbitrariamente un tamaño del árbol y luego se procede a podar el árbol usando la poda de costo-complejidad. El criterio 1 tiene un buen enfoque, pero pierde de vista el hecho que una partición que genera poca ganancia en un nivel n podría generar una alta ganancia en un nivel $n + 1$ o en cualquier otro nivel subsecuente.

El criterio 2 invita a la definición de la Poda Costo-Complejidad para lo cual se define un subárbol $T \subset T_0$ donde T se construye a partir de nodos no terminales de T_0 y sea $|T|$ el número de nodos terminales de T , así mismo se define mediante Fórmula 2. 8 y Fórmula 2. 9

$$N_m = \{x_i \in R_m\}$$

Fórmula 2. 8: Cardinalidad de la región.

$$\hat{c}_m = \frac{1}{N_m} \sum_{x_i \in R_m} y_i$$

Fórmula 2. 9: Definición de \hat{c}_m

Ahora si finalmente se define el criterio de Costo-Complejidad mediante la
Fórmula 2. 10

$$C_\alpha(T) = \sum_{m=1}^{|T|} N_m Q_m(T) + \alpha|T|$$

Fórmula 2. 10: Criterio Costo-Complejidad.

El concepto es encontrar para cada α el subárbol $T_\alpha \subseteq T_0$ tal que minimice $C_\alpha(T)$, es de notar que si α es grande se obtendrá árboles más pequeños, pero si toma su valor mínimo es decir $\alpha = 0$ se obtendrá el mismo árbol T_0 .

Para cada valor de α se ha demostrado que existe un único subárbol más pequeño T_α que logra minimizar $C_\alpha(T)$. Para encontrar T_α se usa la poda del árbol más débil, es decir sucesivamente se contrae el nodo interno que produce el menor aumento por nodo en $\sum_m N_m Q_m(T)$ y se sigue hasta generar el árbol de un único nodo. La estimación de α finalmente se realiza mediante validación cruzada unas 5 a 10 veces y se elige el valor de α que minimiza la suma de cuadrados. El árbol objetivo entonces será T_α .

En la ilustración 2 se pudo observar un ejemplo de la ejecución de un algoritmo de árbol de decisión de clasificación para el caso particular de dos variables x_1 y x_2 continuas y la variable respuesta y de clasificación de color. La variable inicial de

clasificación es x_2 y particionamiento 0.302548, visualmente se aprecia en Ilustración

2. 2

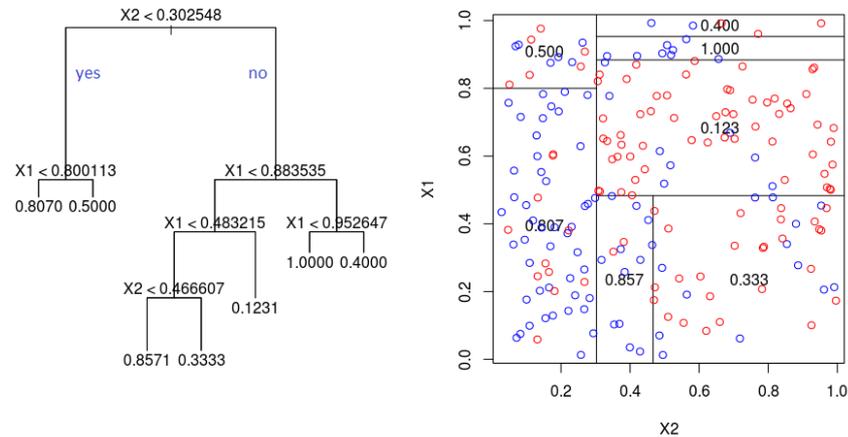


Ilustración 2. 2: Algoritmo Árbol de Decisión.

Algoritmo Q-Learning

Para (Sutton & Barto, 2018) el algoritmo Q-Learning estima la función de valor de estado-acción mediante la Fórmula 2. 11:

$$Q^\pi(S_t, A_t) = Q^\pi(S_t, A_t) + \alpha [R_{t+1} + \gamma Q^\pi(S_{t+1}, A') - Q^\pi(S_t, A_t)]$$

Fórmula 2. 11: Algoritmo Q-Learning

Donde $Q^\pi(S_t, A_t)$ es la función de valor de la política π al aplicar la política A en el estado S en el período t y R_{t+1} es el refuerzo por aplicar la política π estando en el período S_t .

También se usa un parámetro $\alpha \in [0, 1]$ denominado tasa de aprendizaje y representa el grado de actualización de la función valor, a medida que α se acerca a su cota superior 1 el valor de actualización es mayor, en cambio al acercarse α a la

cota inferior 0 es menor el valor de actualización. Este valor de α se multiplica por la diferencia temporal.

La diferencia temporal es el refuerzo obtenido al tomar la acción A en el estado S_t más una ponderación mediante el parámetro γ de $Q^\pi(S_{t+1}, A')$ que es la función de valor de aplicar la política π al aplicar la política A' en el estado S_{t+1} y luego restando $Q^\pi(S_t, A_t)$. El parámetro γ tasa de descuento que se desea dar a la función de valor futura.

Como es de notar la suma de la recompensa y el valor esperado del nuevo estado ponderado se compara con la función de valor actual, si esta diferencia es pequeña el valor actualizado de la función de valor actual no variará significativamente, pero si esta diferencia es alta entonces la función de valor actual se actualizará significativamente, de esta manera al ejecutar el algoritmo de manera infinita, en este caso sería un alto número de iteraciones se garantiza que encontrar una función de valor óptima Q^* .

CAPÍTULO 3

3. Metodología

3.1. Fuente de Datos

La empresa “XYZ” se encuentra domiciliada en la ciudad de Guayaquil, misma que se dedica a la producción de alimentos balanceados de mascotas. Para empresas de producción de este tipo es de vital importancia la optimización de los procesos con la finalidad de reducir tiempos y costos de producción.

Se considerarán como unidades de investigación las órdenes de producción de alimentos balanceados de mascotas desde el 2015 al 2018, considerando que la empresa labora 24 horas al día distribuido en 2 turnos de 12 horas cada uno (07h00am-06h59pm diurno y 07h00pm-06h59am nocturno) los 365 días del año.

La base de datos está compuesta por 17 variables y 22405 observaciones.

3.2. Software

Los diferentes algoritmos que se usarán serán implementados en el lenguaje de programación R (Ihaka R, Gentleman R. (1993). R. <https://www.r-project.org/>) mediante el entorno de desarrollo integrado R Studio (Allaire J. (2011). R Studio. <https://www.rstudio.com/>) con los siguientes paquetes:

- readxl.
- lubridate.
- tidyverse.
- modeest.

- knitr.
- moments.
- e1071.
- ggplot2.
- rpart.
- caret

3.3. Metodología de Trabajo

Se procederá a implementar la metodología propuesta por (Wickham & Golemund, 2017) en su libro “R for Data Science” en el cual detalla las etapas para un proyecto de ciencia de datos bajo mediante Ilustración 3. 1

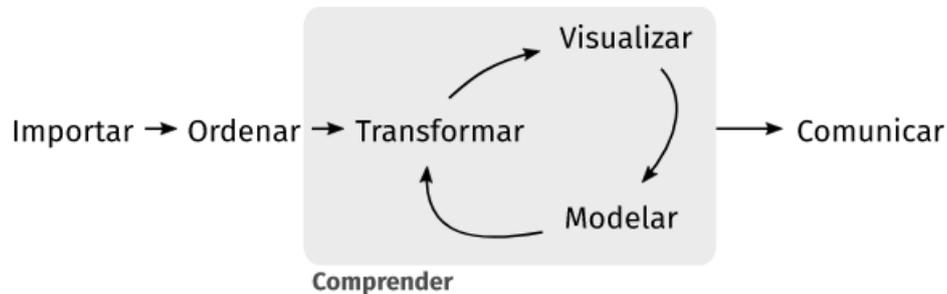


Ilustración 3. 1: Etapas de Proyecto

3.4. Etapa de Importar.

En esta etapa se procede a importar la base de datos a R Studio a partir de un archivo Excel mediante el paquete readxl.

3.5. Etapa Ordenar

Posterior a la importación de datos se convirtió en data frame para mejor manejo, teniendo 17 variables y 22405 registros, estas se detallan en Tabla 3. 1

No	Variable	Descripción	Tipo Variable
01	NO_ORDEN	Número de Orden de Producción.	Categórica
02	N_SUB_ORDEN	Número de Suborden de Producción.	Categórica
03	PRESENTACION	Presentación del producto, puede ser; E: Ensacado, G: Granel, R: Reproceso.	Categórica
04	SKU	Código único del producto.	Categórica
05	HORA_INICIO	Hora de inicio de orden de producción.	Cuantitativa
06	HORA_FIN	Hora de fin de orden de producción.	Cuantitativa
07	HORA_DIF_INI_FIN	Diferencia entre la hora de inicio y fin de la hora de producción en horas, minutos y segundos.	Cuantitativa
08	DIF_HORA	Tiempo de producción en horas.	Cuantitativa
09	KG	Kilogramos producidos en la orden de producción.	Cuantitativa
10	LOTE	Número de lote asignado a la orden de producción.	Categórica
11	RAP_TON_HORA	Rapidez de toneladas producidas por hora.	Cuantitativa
12	CAP_USA_PON	Capacidad Usada Ponderada.	Cuantitativa
13	RAP_KG_HORA	Rapidez de kilogramos producidos por hora.	Cuantitativa
14	T_MEZCLA	Grado de mezcla.	Cuantitativa
15	T_DOSIS	Grado de dosis.	Cuantitativa
16	FECHA_HORA_ENC	Fecha y hora de encendido de la máquina.	Cuantitativa
17	TIPO_PROCESO	Tipo de proceso, puede ser: Pellet, Expandido, Polvo o Null.	Categórica

Tabla 3. 1: Variables iniciales

3.6. Etapa Transformar

Una vez importados y ordenados los datos se procede a transformarlos mediante las siguientes subetapas:

- Reducción de variables.
- Creación de variables.
- Estadísticos Descriptivos.

En la subetapa de reducción de variables se analizan las variables que no aportan información relevante para el análisis, las variables excluidas se detallan en

Tabla 3. 2

No	Variable	Motivo Exclusión
01	HORA_DIF_INI_FIN	Se creó la variable SEP_UP a partir de esta variable.
02	RAP_KG_HORA	Redundante con la variable rapidez toneladas/hora
03	CAP_USA_PON	Variable que mide la capacidad de la planta.
04	T_MEZCLA,	Variable del departamento de calidad.
05	T_DOSIS	Variable del departamento de calidad.
06	FECHA_HORA_ENC	Se crearon 2 variables nueva a partir de esta

Tabla 3. 2: Variables excluidas.

En la subetapa de creación de variables se crean aquellas variables que se consideran de interés relevante para la investigación, se han creado nuevas variables a partir de variables ya existentes, estas se detallan en Tabla 3. 3:

No	Variable	Descripción	Tipo Variable
01	FECHA_ENC	Fecha de encendido de la máquina.	Cuantitativa
02	HORA_ENC	Hora de encendido de la máquina.	Cuantitativa
03	TIME_PRE	Tiempo de preparación en horas.	Cuantitativa
04	DIA	Día de la semana de la orden de producción.	Categórica
05	MES	Mes de la orden de producción.	Categórica
06	ANIO	Año de la orden de producción.	Categórica
07	TURNO_INI	Turno de inicio de la orden de producción	Categórica

Tabla 3. 3: Variables creadas.

Considerando las nuevas variables y la exclusión de otras finalmente se cuenta con las variables detalladas en Tabla 3. 4

No	Variable	Descripción	Tipo Variable
01	NO_ORDEN	Número de Orden de Producción.	Categórica
02	N_SUB_ORDEN	Número de Suborden de Producción.	Categórica
03	PRESENTACION	Presentación del producto, puede ser; E: Ensacado, G: Granel, R: Reproceso.	Categórica
04	SKU	Código único del producto.	Categórica
05	HORA_INICIO	Hora de inicio de orden de producción.	Cuantitativa
06	HORA_FIN	Hora de fin de orden de producción.	Cuantitativa
07	DIF_HORA	Tiempo de producción en horas.	Cuantitativa

08	KG	Kilogramos producidos en la orden de producción.	Cuantitativa
09	LOTE	Número de lote asignado a la orden de producción.	Categórica
10	RAP_TON_HORA	Rapidez de toneladas producidas por hora.	Cuantitativa
11	TIPO_PROCESO	Tipo de proceso, puede ser: Pellet, Expandido, Polvo o Null.	Categórica
12	FECHA_ENC	Fecha de encendido de la máquina.	Cuantitativa
13	HORA_ENC	Hora de encendido de la máquina.	Cuantitativa
14	TIME_PRE	Tiempo de preparación.	Cuantitativa
15	DIA	Día de la semana de inicio de la orden de producción	Categórica
16	MES	Mes de la orden de producción	Categórica
17	ANIO	Año de la orden de producción.	Categórica
18	TURNO_INI	Turno de inicio de la orden de producción.	Categórica

Tabla 3. 4: Variables finales.

Finalmente, en la subetapa de estadísticos se procedió a realizar un análisis descriptivo de las variables cuantitativas tiempo de producción, kilogramos, rapidez y tiempo de preparación, pero segregado por cada una de las siguientes variables categóricas:

- **Presentación:** ensacado, granel y reproceso.
- **Tipo Proceso:** polvo, pellet y expandido.
- **Turno:** diurno y nocturno.
- **Día:** lunes, martes, miércoles, jueves, viernes, sábado y domingo.
- **Mes:** enero, febrero, marzo, abril, mayo, junio, julio, agosto, septiembre, octubre, noviembre y diciembre.

- **Año:** 2015, 2016, 2017, 2018.

De las variables cuantitativas segregadas por variables categóricas previamente descritas se extraen las medidas de tendencia central y dispersión.

3.7. Etapa Visualización

Esta etapa de visualización es complementaria a la subetapa de estadísticos de la etapa de transformación ya que permite observar el comportamiento espacial de las variables.

De las variables categóricas presentación, tipo de proceso, turno, día, mes y año se extraen gráficos de barra, respecto a las variables cuantitativas se extraen histogramas, adicionalmente se grafica una distribución normal para saber si tiene un comportamiento normal.

Al contarse con variables cuantitativas y categóricas es de interés buscar patrones en su comportamiento de las variables cuantitativa por cada variable categórica, con esta finalidad se construye un diagrama de cajas para cada variable cuantitativa pero segregado por cada variable categórica, esto se lo realiza mediante el paquete ggplot2.

3.8. Etapa de Modelización.

3.8.1. Estimación de Funciones de Distribución de Probabilidad

Una vez analizado el comportamiento estadístico, así como realizada la visualización de datos se infiere que es importante obtener las funciones de distribución de ciertas variables aleatorias implícitas en la secuencia de producción, estas distribuciones son:

- X_1 : Variable aleatoria de los tiempos de para entre la producción de 2 SKU's.
- X_2 : Variable aleatoria de la rapidez de producción de los SKU's.

Una vez que se conocen las variables aleatorias $X_1, y X_2$ se estima sus funciones de distribución de probabilidad f_1, f_2 . Para la estimación de las funciones de distribuciones de probabilidad se usará el paquete univariateML, se probarán estimarán las distribuciones:

- Gamma.
- Log Gamma.
- Gamma Inversa.
- Beta.
- Exponencial.
- Normal.
- Log Normal.
- Gausiana Inversa.

- Weibull.
- Weibull Inversa.
- Rayleigh.

Luego de estimadas las distribuciones se procede a compararlas, los criterios de comparación son:

- Criterio de Información de Akaike: $AIC = 2k - 2 \ln(L)$
- Criterio de Información Bayesiano: $BIC = k \ln(n) - 2 \ln(L)$

Donde:

- **k**: Número de parámetros de la distribución.
- **L**: Máxima probabilidad de la función de verosimilitud de la distribución.
- **n**: Número de observaciones.

Se considera que a menor valor de AIC o BIC el modelo es de mejor calidad que otro de mayor AIC o BIC.

3.8.2. Algoritmo Generador de Secuencias

Adicionalmente a las variables previamente descritas es de importancia considerar que en el proceso de producción existen productos que por su propia naturaleza no se pueden producir de manera secuencial ya que generan contaminación cruzada, por este motivo es necesario realizar una limpieza de las máquinas, los tiempos de limpieza son conocidos por el departamento de calidad y se la modeliza como una matriz de estados según el producto producido inicialmente y el siguiente, la matriz se detalla en Tabla 3. 5:

$$M_{(i,j)} = \begin{pmatrix} Gr/Gr & EN & PV & PL & PO & CD & CM & FS & FG & GD & CB & CY \\ EN & 0 & 1 & 1 & 1 & 0 & 2 & 2 & 2 & 0 & 2 & 1 \\ PV & 0 & 0 & 1 & 1 & 0 & 1 & 2 & 2 & 0 & 2 & 1 \\ PL & 0 & 0 & 0 & 0 & 0 & 1 & 2 & 2 & 0 & 2 & 0 \\ PO & 0 & 0 & 0 & 0 & 0 & 1 & 2 & 2 & 0 & 2 & 0 \\ CD & 0 & 0 & 1 & 1 & 0 & 0 & 2 & 0 & 0 & 2 & 1 \\ CM & 2 & 2 & 1 & 1 & 0 & 0 & 2 & 0 & 2 & 2 & 2 \\ FS & 2 & 2 & 1 & 1 & 1 & 1 & 0 & 0 & 2 & 2 & 2 \\ FG & 2 & 2 & 2 & 2 & 0 & 0 & 2 & 0 & 2 & 2 & 2 \\ GD & 0 & 1 & 1 & 1 & 0 & 1 & 2 & 2 & 0 & 2 & 1 \\ CB & 0 & 0 & 0 & 1 & 0 & 1 & 2 & 0 & 0 & 0 & 0 \\ CY & 0 & 1 & 0 & 1 & 0 & 1 & 2 & 2 & 0 & 2 & 0 \end{pmatrix}$$

Tabla 3. 5: Matriz de Estados.

Donde $EN, PV, PL, PO, CD, CM, FS, FG, GD, CB, CY$ son los grupos de productos, cada producto pertenece a un único grupo, adicionalmente los estados 0, 1 y 2 significan 0, 12 y 70 minutos de limpieza y acondicionamiento.

Esta matriz permite definir una variable determinística de los tiempos de limpieza entre dos SKU's, es decir:

$X_{(i,j)}$: Tiempo de limpieza al producir el SKU i y luego el SKU j .

Una vez consideraras las variables estocásticas y determinística a usar se implementa un algoritmo que recibe como input el Plan Maestro de Producción que son las órdenes de producción de los SKU a producir, es decir:

- Fecha.
- Turno.
- SKU's a producir.
- Presentación.

- Kg.

Luego el usuario elige el SKU con el cuál iniciar el proceso de producción, esta decisión es tomada por el usuario de manera empírica según su experiencia, una vez que se termina de producir este SKU inicial se toman los siguientes datos:

- SKU producido.
- Hora de inicio del proceso de producción de SKU.
- Hora de finalización del proceso de producción del SKU.

Luego el algoritmo recibe los valores indicados previamente y actualiza la función de densidad f_2 , este es el reforzamiento del ambiente que permite absorber la información de la rapidez de producción del SKU. También se actualiza el tiempo de producción acumulado según Fórmula 3. 1:

$$Tiempo_{Acumulado} < -Tiempo_{Acumulado} + (Hora_{ini} - Hora_{fin})$$

Fórmula 3. 1: Actualización tiempo acumulado con SKU inicial.

El siguiente paso es sugerir un nuevo SKU de los productos del Plan Maestro de Producción, se definen 2 criterios para la elección del siguiente SKU:

Criterio 1: El SKU j candidato minimiza el tiempo de para entre la producción el SKU inicial y el candidato, es decir se toma el valor de la función de distribución f_1 para la secuencia SKU i y luego el SKU j .

Criterio 2: El SKU j candidato es el de máxima rapidez de producción, es decir se toma el valor de la función de distribución f_2 .

Para ambos criterios una vez sugerido el nuevo SKU e iniciado el proceso de producción se toma un nuevo valor que es el tiempo real de para entre la finalización de la producción del SKU inicial y el inicio de la producción del SKU elegido, este tiempo de para actualiza el tiempo de producción acumulado según Fórmula 3. 2:

$$Tiempo_{Acumulado} < -Tiempo_{Acumulado} + (Hora_{Fin_{SKU(i)}} - Hora_{Ini_{SKU(i')}})$$

Fórmula 3. 2: Actualización tiempo acumulado con tiempo de para.

Donde i es el SKU inicial y i' es el SKU que sigue en la secuencia de producción.

Finalmente se obtienen 2 secuencias, una por cada criterio, ambas cumplirán con el plan de maestro de producción, en el caso particular de iniciar de un Plan Maestro de Producción con los productos: $SKU_1, SKU_2, SKU_3, SKU_4$, las secuencias serán:

$$Secuencia_1: SKU_1, \quad SKU_3, \quad SKU_2, \quad SKU_4.$$

$$Secuencia_2: SKU_1, \quad SKU_4, \quad SKU_3, \quad SKU_2.$$

Ahora se debe elegir una secuencia, para esto se las clasificará como “Secuencia Óptima” y “Secuencia No Óptima”, el criterio de optimalidad será la de mínimo tiempo de para de la secuencia según Fórmula 3. 3:

$$Secuencia^* = \min\{Tiempo\ Para_{Secuencia_1}, Tiempo\ Para_{Secuencia_2}\}$$

Fórmula 3. 3: Elección Secuencia Óptima.

Este criterio permite conocer aquellas secuencias en la cual el criterio 1 es mejor que el criterio 2 y viceversa, para etiquetar la secuencia se define una variable dicotómica llamada etiqueta según Fórmula 3. 4:

$$etiqueta_i \begin{cases} 1; & \text{Si la secuencia } i \text{ es óptima} \\ -1; & \text{Si la secuencia } i \text{ no es óptima.} \end{cases}$$

Fórmula 3. 4: Variable etiqueta.

3.8.2.1. Conjunto 1: Variables Predictoras: Indicadoras, Categóricas y Cuantitativas

Se construye un conjunto de variables predictoras, agrupadas en 3 tipos de variables:

- Variables Indicadoras: Se crean 210 variables indicadoras, una por cada SKU del catálogo de productos fabricados según Fórmula 3. 5:

$$\circ \quad SKU_i = \begin{cases} 1; & \text{Si el SKU } i \text{ forma parte de la secuencia.} \\ 0; & \text{Si el SKU } i \text{ no forma parte de la secuencia.} \end{cases}$$

Fórmula 3. 5: Variables indicadoras.

- Variables categóricas: Se crean 2 variables categóricas, donde:
 - $DIA_SEMANA_i =$ Día de la semana de la secuencia i .
 - $TURNO_i =$ Turno de la secuencia i .
- Variables cuantitativas: Se crean 4 variables cuantitativas, donde:
 - $TIEMPO_PROD_i:$ Horas tomadas para producir la secuencia i .
 - $TIEMPO_PARA_i:$ Horas de para en la producción se secuencia i .
 - $RAPIDEZ_i:$ Tonelas producidas entre Horas tomadas en la secuencia i .

- $TONELADAS_i$: Toneladas producidas en lse secuencia i .

3.8.2.2. Conjunto 2: Variables Predictoras: Categóricas y Cuantitativas.

También se construirá un segundo conjunto de variables predictoras considerando las variables categóricas y cuantitativas del Conjunto 1, pero excluyendo las variables indicadoras con la finalidad de analizar la sensibilidad de los algoritmos a estas variables.

3.8.3. Algoritmos de Clasificación

Se ejecutará el algoritmo de generación de secuencias para 2372 planes de producción, de esta manera se obtendrá suficientes secuencias etiquetadas, estas serán el input para un algoritmo de clasificación que dado un plan maestro de producción pueda generar las secuencias candidatas y clasificarlas como óptima o no óptima.

Para el aprendizaje estadístico de clasificación se procederá a analizar dos algoritmos de clasificación como lo son:

- Máquinas Vector Soporte.
- Árbol de Clasificación.

3.8.3.1. Máquinas Vector Soporte

Según (James, Witten, Hastie, & Tibshirami, 2013) el algoritmo de máquina vector soporte es un caso particular del Clasificador de Vector Soporte donde para un conjunto de variables $\{x_1, x_2, \dots, x_p\}$ de dimensión p se define un hiperplano de dimensión $p - 1$, además se tiene una variable y denominada respuesta donde los valores de esta variable se denominan clases.

El clasificador de vector soporte necesita del cálculo del producto interno de las observaciones según Fórmula 3. 6:

$$\langle x_i, x_{i'} \rangle = \sum_{j=1}^p x_{ij} x_{i'j}$$

Fórmula 3. 6: Producto Interno Vectorial

Este producto interno se lo puede generalizar de la forma $K(x_i, x_{i'})$ donde K es una función denominada kernel y cuantifica la diferencia entre dos observaciones, esta cuantificación se realiza mediante la correlación de Pearson. Finalmente, esta función kernel se puede definir según Fórmula 3. 7 y Fórmula 3. 8:

$$K(x_i, x_{i'}) = \left(1 + \sum_{j=1}^p x_{ij} x_{i'j}\right)^d$$

Fórmula 3. 7: Kernel Polinomial

Donde este kernel se denomina Kernel Polinomial de grado d entero positivo.

$$K(x_i, x_{i'}) = \exp(-\gamma \sum_{j=1}^p (x_{ij} - x_{i'j})^2)$$

Fórmula 3. 8: Kernel Radial

Donde este kernel se denomina Kernel Radial con parámetro γ positivo.

Para ambos modelos es necesaria la definición de los hiperparámetros C y γ para encontrar el modelo de mejor ajuste.

3.8.3.2. Árbol de Clasificación

Para (James, Witten, Hastie, & Tibshirami, 2013) el algoritmo de clasificación consiste en dos etapas:

- **Etapas 1:** División del espacio en regiones R_1, R_2, \dots, R_j .
- **Etapas 2:** Predecir la variable respuesta y en cada región.

Se debe decidir en qué predictores y en qué valores de estos se realiza la partición, esta partición se denomina nodo, a medida que se sigue dividiendo el espacio en regiones se genera un árbol con diferentes ramas.

Para cumplir con las etapas se aplica el algoritmo de División Binaria Recursiva, siguiendo los pasos:

- **Paso 1:** Se inicia con todas las observaciones perteneciendo a una única región R .
- **Paso 2:** Para las variables categóricas X_j el umbral de división s es el nivel de esta, para las variables cuantitativas X_j se ordenan sus valores

de menor a mayor y se toma como un umbral s la semisuma de cada par de valores.

- **Paso 3:** Para las 2 regiones generadas se calcula una métrica, luego se suma el valor de estas dos medidas para todas las regiones generadas en el Paso 2 y por cada variable predictora X_j , es decir al final de este paso se obtiene un único valor por cada variable predictora.

Existen 3 métricas que se pueden considerar:

- Tasa de Error de Clasificación Fórmula 3. 9:

$$E_m = 1 - \max_k(\hat{p}_{mk})$$

Fórmula 3. 9: Métrica Tasa de Error de Clasificación.

- Índice Gini Fórmula 3. 10:

$$G_m = \sum_{k=1}^K \hat{p}_{mk} (1 - \hat{p}_{mk})$$

Fórmula 3. 10: Métrica Índice Gini

- Entropía Cruzada Fórmula 3. 11:

$$D = - \sum_{k=1}^K \hat{p}_{mk} \log(\hat{p}_{mk})$$

Fórmula 3. 11: Métrica Entropía Cruzada

Donde \hat{p}_{mk} es la proporción de observaciones que pertenecen a la clase menos frecuente.

- **Paso 4:** Se selecciona la variable predictora X_j que tiene la menor medida.
- **Paso 5:** Se repiten los pasos 1, 2, 3 y 4 iterativamente para las regiones generadas en el paso anterior hasta que se cumpla una norma de para, esta norma puede ser que el árbol tenga un máximo de nodos terminales, entre otros.

Cada uno de los umbrales que dividen las regiones se denominan nodos.

3.8.3.3. Comparación de Algoritmos de clasificación.

Inicialmente se hizo una partición del conjunto de datos, en dos conjuntos disjuntos, estos serán:

- Conjunto de entrenamiento: 80% de las observaciones.
- Conjunto de prueba: 20% de las observaciones.

Luego se ejecutarán los 2 algoritmos:

Máquina Vector Soporte:

Se procederá a generar valores para los hiperparámetros C y γ tanto para el modelo con kernel radial, así como para el kernel polinomial, luego se elegirá el valor de C y γ tal que estos generen el mínimo error.

Árbol de Clasificación.

Al contarse con 217 variables para el conjunto 1, la profundidad del árbol dada por el parámetro cp será de 0.01 teniendo en consideración que este es un parámetro que pertenece al intervalo $[0,1]$ y que a medida que el parámetro se acerca a 0 la profundidad del árbol será máxima y que $cp=1$ no genera divisiones.

Se evaluará posteriormente el valor del Deviance, mismo que está dado por Fórmula 3. 12:

$$Deviance = -2 \sum_m \sum_k n_{mk} \log(\hat{p}_{mk})$$

Fórmula 3. 12: Deviance

A menor valor del Deviance se considera que es un mejor ajuste del modelo.

Una vez ejecutados los 2 algoritmos se procederá a medir el rendimiento de los algoritmos mediante la matriz de confusión mediante el uso del paquete caret, a partir de la matriz de confusión Tabla 3. 6:

		Clase Verdadera	
		Caso Positivo	Caso Negativo
Clase Hipotética	Caso Positivo	A	B
	Caso Negativo	C	D

Tabla 3. 6: Matriz de Confusión

- **A:** Verdadero Positivo.
- **B:** Falso Positivo.
- **C:** Falso Negativo.
- **D:** Verdadero Negativo.

Con esta matriz se calcularán las siguientes métricas Tabla 3. 7:

Métrica	Calculo	Significado
Exactitud	$\frac{A + D}{A + B + C + D}$	Proporción de predicciones correctamente clasificados.
Precisión	$\frac{A}{A + B}$	Proporción de predicciones positivas correctamente clasificadas.
Sensibilidad	$\frac{A}{A + C}$	Proporción de casos positivos correctamente clasificados.
Especificidad	$\frac{D}{B + D}$	Proporción de casos negativos correctamente clasificados.
Prevalencia	$\frac{A + C}{A + B + C + D}$	Proporción de casos positivos.
Tasa de Retención	$\frac{A}{A + B + C + D}$	Proporción total de casos positivos correctamente clasificados.
Retención de Prevalencia	$\frac{A + B}{A + B + C + D}$	Proporción de predicciones positivas.
Exactitud Balanceada	$\frac{\frac{A}{A + C} + \frac{D}{B + D}}{2}$	Semisuma de la sensibilidad y especificidad.

Tabla 3. 7: Métricas

Adicionalmente se medirá el estadístico Kappa, este estadístico ajusta la exactitud, pero considerando el hecho que se puede clasificar correctamente como producto del azar y no como producto de la fiabilidad del algoritmo.

Este estadístico se estima mediante Fórmula 3. 13:

$$\kappa = \frac{\text{Pr}(a) - \text{Pr}(e)}{1 - \text{Pr}(e)}$$

Fórmula 3. 13: Estadístico Kappa.

Donde $\text{Pr}(a)$ es la proporción del acuerdo real y $\text{Pr}(e)$ es el acuerdo esperado entre el clasificador y los valores reales, considerando el hecho que fueron elegidos al azar. Kappa es un estadístico pertenece al intervalo $[0,1]$, donde se espera que el valor sea lo más cercano a 1, donde 1 representa la máxima concordancia y 0 representa nula concordancia.

Los valores de κ se interpretan según Tabla 3. 8:

Concordancia	κ
Pobre	<0.20
Regular	0.20-0.40
Moderada	0.40-0.60
Buena	0.60-0.80
Muy Buena	>0.8

Tabla 3. 8: Interpretación Estadístico Kappa

También se extraerán los siguientes estadísticos para realizar inferencias

Tabla 3. 9:

Estadístico	Interpretación
95% IC	Intervalo de confianza del 95% para la exactitud
Tasa sin Información	Mayor proporción de las clases observadas
Valor-p [Acc>NIR]	$H_0: \text{Acc} < \text{NIR}$ $H_1: \text{Acc} > \text{NIR}$
Mcnemas's Test	$H_0: \text{Frecuencia Marginal Fila} = \text{Frecuencia Marginal Columna}$
Valor-p	$H_1: \text{Frecuencia Marginal Fila} = \text{Frecuencia Marginal Columna}$

Tabla 3. 9: Estadísticos

Finalmente dadas las métricas se comparará el desempeño de los algoritmos, luego se procederá a evaluar los datos de prueba en todos los modelos.

CAPÍTULO 4

4. Resultados

4.1. Análisis Exploratorio Descriptivo

Se realizaron los siguientes análisis:

- Univariante:
 - Cuantitativo.
 - Categórico.
- Bivariante:
 - Univariante Cuantitativo vs Multivariante Categórico.
 - Bivariante Cuantitativo vs Multivariante Categórico.

4.1.1. Análisis Univariante Cuantitativo.

Se procede a extraer las medidas de tendencia central y medidas de dispersión de las variables cuantitativas: tiempo de producción, rapidez, kilogramos y tiempo de preparación, pero desglosadas por cada una de las variables categóricas: Presentación, Tipo de Proceso, Turno, Día, Mes, Año obteniéndose 7.1

Se puede analizar de acuerdo a las medidas de tendencia central y dispersión descritas que las variables tienen un comportamiento muy similar independientemente de las distintas variables categóricas por las cuáles se segregó el conjunto de datos, esto era esperable puesto que la planta de producción labora todos los días de la semana, las 24 horas del día y los diferentes procesos están estandarizados.

Finalmente se procede a generar el histograma de las variables cuantitativas tiempo de producción, kilogramos, rapidez y tiempo de para, obteniéndose Ilustración 4: 1, Ilustración 4: 2, Ilustración 4: 3 e Ilustración 4: 4

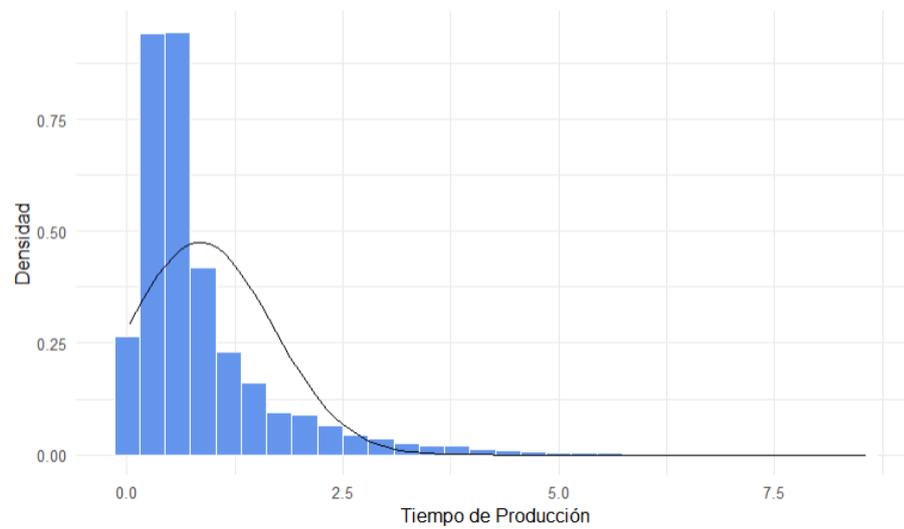


Ilustración 4: 1: Histograma Tiempo de Producción

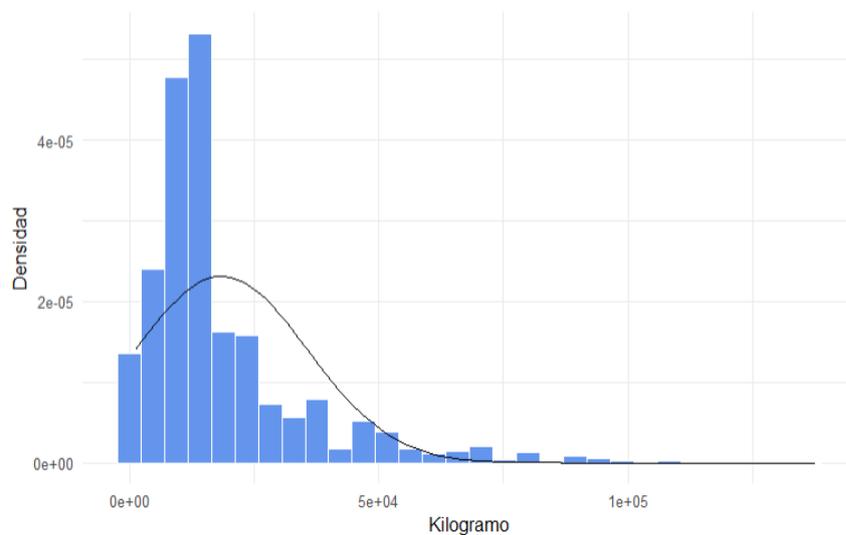


Ilustración 4: 2: Histograma de Kilogramos.

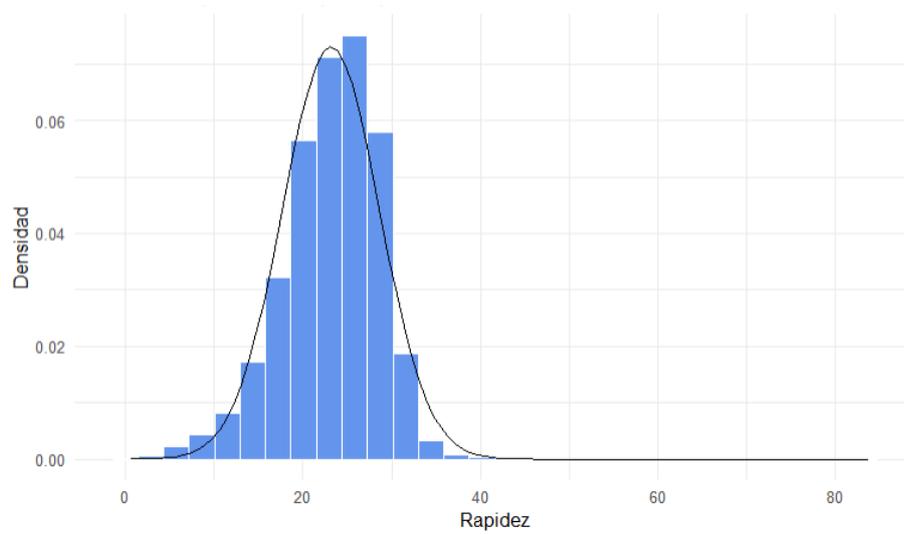


Ilustración 4: 3: Histograma Rapidez.

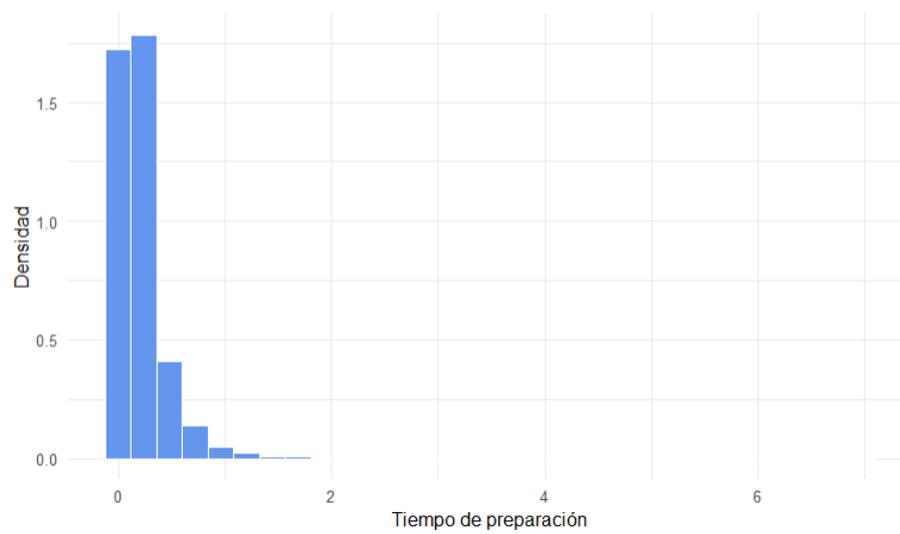


Ilustración 4: 4: Histograma Tiempo de preparación.

Se puede observar que la distribución de las variables tiempo de producción, kilogramos y tiempo de preparación tienen un sesgo positivo, en cambio la distribución de la rapidez es aproximadamente insesgada.

4.1.2. Análisis Univariante Categórico.

Para las variables categóricas Tipo de proceso, turno, día, mes y año se procederá a generar un gráfico de barras, obteniéndose 7.2

Presentación

Se puede apreciar que la mayoría de las órdenes de producción corresponden a productos de presentación ensacado, seguido de granel y reproceso.

Tipo de Proceso

EL tipo de proceso pellet y polvo son más frecuentes, luego expandido y null.

Turno

Al trabajar la empresa 24/7, las órdenes de producción procesadas están distribuidas aproximadamente igual en los 2 turnos, en este caso la cantidad de órdenes de producción procesadas en el turno diurno supera al turno nocturno.

Día

Al igual que con los turnos, la distribución de la cantidad de órdenes de producción a lo largo de los días de semana es aproximadamente igual.

Mes

Se aprecia que el mes de mayor cantidad de órdenes de producción procesadas es agosto, luego se puede formar un grupo con junio, julio, septiembre y octubre y un último grupo con los meses restantes.

Año

Se puede apreciar que años con menor cantidad de órdenes procesadas son 2015 y 2018 y que en los años 2016 y 2017 son aproximadamente iguales.

4.1.3. Análisis Bivariante: Univariante Cuantitativo vs Multivariante Categórico

Se procedió a generar un diagrama de cajas de cada variable cuantitativa segregada por cada variable categórica, obteniéndose los resultados 7.3

En todos los gráficos de cajas se puede comprobar la abundancia de datos atípicos, pero en este caso estos no obedecen a errores en la recolección de los datos o de otro origen, estos obedecen a que los volúmenes de producción están directamente influenciados a los pedidos de los clientes, por esto existe una gran variación en todos los análisis.

4.1.4. Análisis Bivariante: Bivariante Cuantitativo vs Multivariante Categórico

Se procedió a generar un gráfico de dispersión bidimensional para todos los pares de variables cuantitativas segregada por cada variable categórica, obteniéndose 7.4

Para las diferentes las variables categóricas evaluadas para los gráficos de dispersión bidimensional de las diferentes variables cuantitativas se puede concluir que no se puede detectar un patrón en el comportamiento de las variables cuantitativas por nivel de las categóricas.

Una vez realizado el análisis exploratorio descriptivo, así como la visualización de los datos se puede concluir que la distribución de las diferentes variables del proceso productivo no se ve influenciada por las variables categóricas analizadas.

4.2. Estimación de Funciones de Distribución de Probabilidad.

Dado 2372 planes de producción se procede a estimar la distribución de las variables aleatorias:

- X_1 : Tiempo de para entre la producción de dos SKU's (Horas)
- X_2 : Rapidez (Toneladas/Hora) en la producción de un SKU.

Primero se procede a extraer el soporte S_1 y S_2 de las variables aleatorias X_1 y X_2 obteniéndose los siguientes resultados:

- S_1 : Se generaron 4704 permutaciones de 2 pares de SKU's, lo que implica que se tendrán 4704 soportes, este valor denota que solo se produce un 10.71% de todas las combinaciones posibles puesto que de los 210 SKU's del catálogo el número de combinaciones en las que se pueden producir 2 productos está dado por $P\binom{210}{2} = 43890$.
- S_2 : Se generaron 210 soportes.

Posteriormente se procede a estimar las funciones de distribución de probabilidad f_1 y f_2 , para esto se comparan las funciones recomendadas según el criterio AIC y BIC obteniéndose que ambos criterios dieron la misma función como la que mejor se ajusta, obteniéndose los siguientes resultados:

- f_1 : A continuación, se detallan 3 ejemplos de funciones ajustadas
Ilustración 4: 5 Ilustración 4: 6 Ilustración 4: 7

$$X \sim \text{G Inv}(\mu = 0.07548, \lambda = 0.64542) \quad f(x) = \sqrt{\frac{\lambda}{2\pi x^3}} \exp\left(-\frac{\lambda(x - \mu)^2}{2\mu^2 x}\right)$$

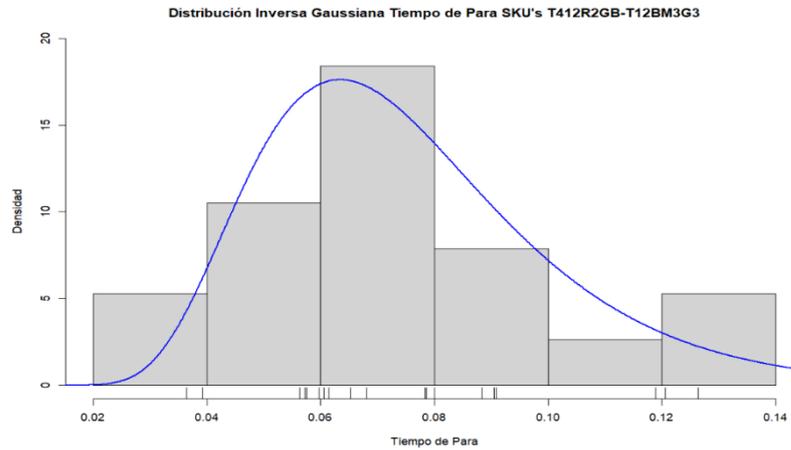


Ilustración 4: 5: Distribución Inversa Gaussiana Tiempo de Para.

$$X \sim \text{Weibull}(\alpha = 1.65397, \lambda = 0.09906) \quad f(x) = \lambda \alpha (\lambda x)^{\alpha-1} e^{-(\lambda x)^\alpha}$$

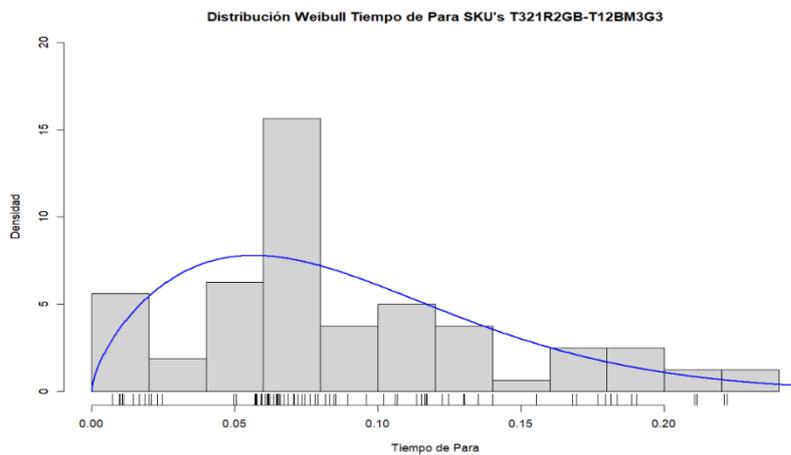


Ilustración 4: 6: Distribución Weibull Tiempo de Para.

$X \sim \text{Exponencial}(\lambda = 1.853) \quad f(x) = \lambda e^{-\lambda x}$

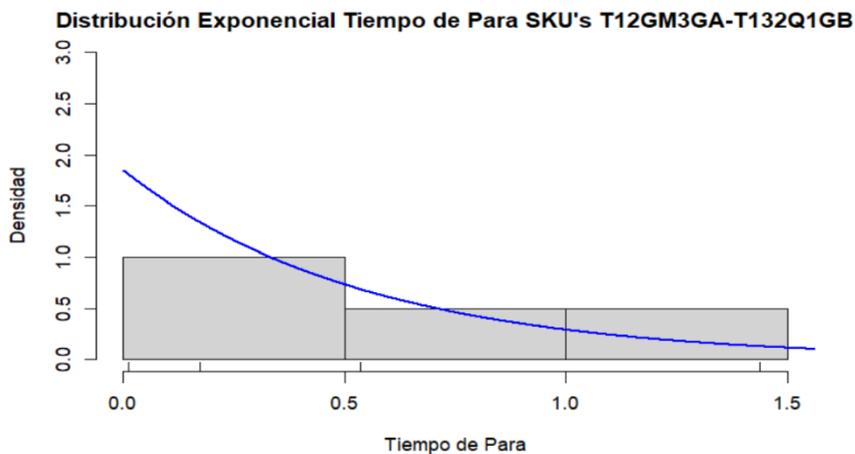


Ilustración 4: 7: Distribución Exponencial Tiempo de Para.

La frecuencia de las distribuciones de tiempo de para está dada por Ilustración

4: 8

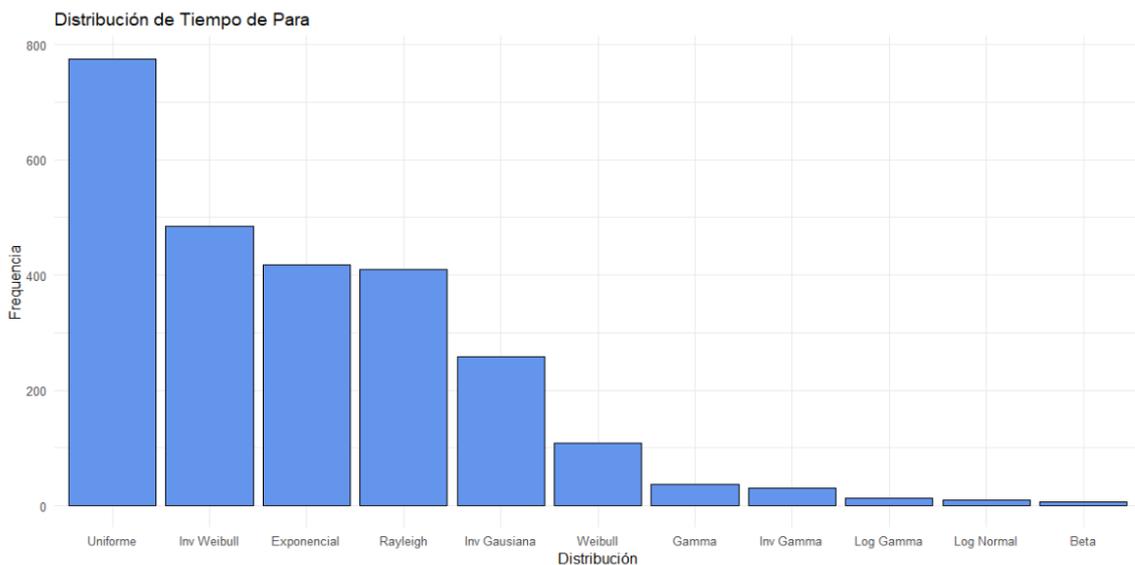


Ilustración 4: 8: Frecuencia de Distribuciones de Tiempo de Para.

- f_2 : A continuación, se detallan 3 ejemplos de funciones ajustadas

Ilustración 4: 9 Ilustración 4: 10 Ilustración 4: 11

$$X \sim \text{Weibull}(\alpha = 8.318, \lambda = 22.906) \quad f(x) = \lambda \alpha (\lambda x)^{\alpha-1} e^{-(\lambda x)^\alpha}$$

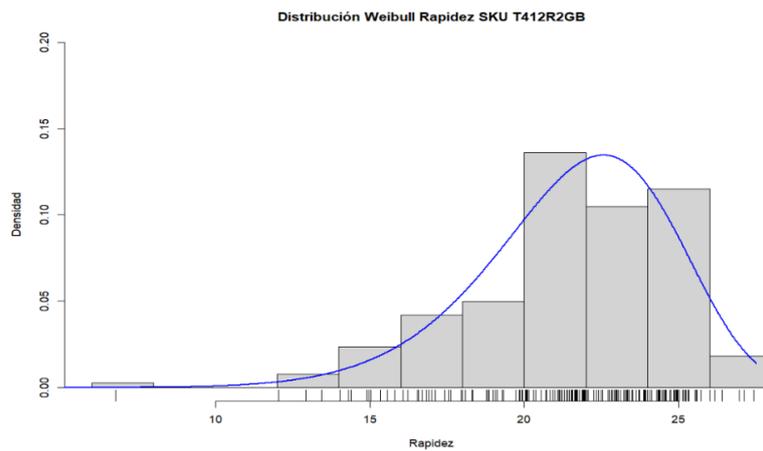


Ilustración 4: 9: Distribución Weibull Rapidez.

$$X \sim \text{G Inv}(\alpha = 30.21, \beta = 864.95) \quad f(x) = \frac{\beta^\alpha}{\Gamma(\alpha)} (1/x)^{\alpha+1} \exp(-\beta/x)$$

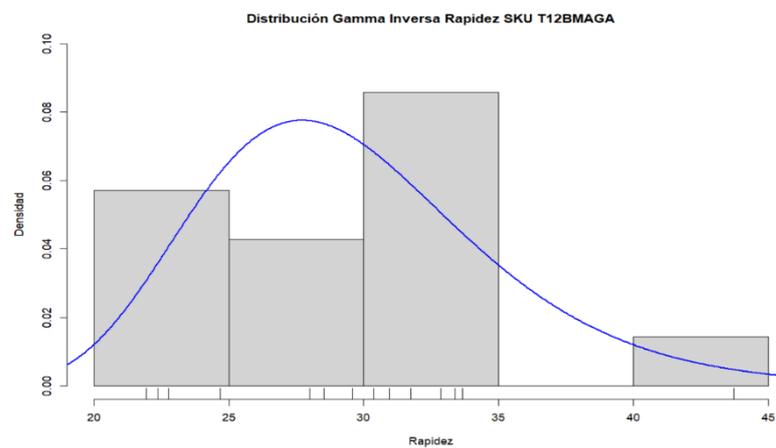


Ilustración 4: 10: Distribución Gamma Inversa Rapidez.

$$X \sim \text{Gamma}(\alpha = 28.125, \lambda = 1.119) \quad f(x) = \frac{\lambda(\lambda x)^{\alpha-1} e^{-\lambda x}}{\Gamma(\alpha)}$$

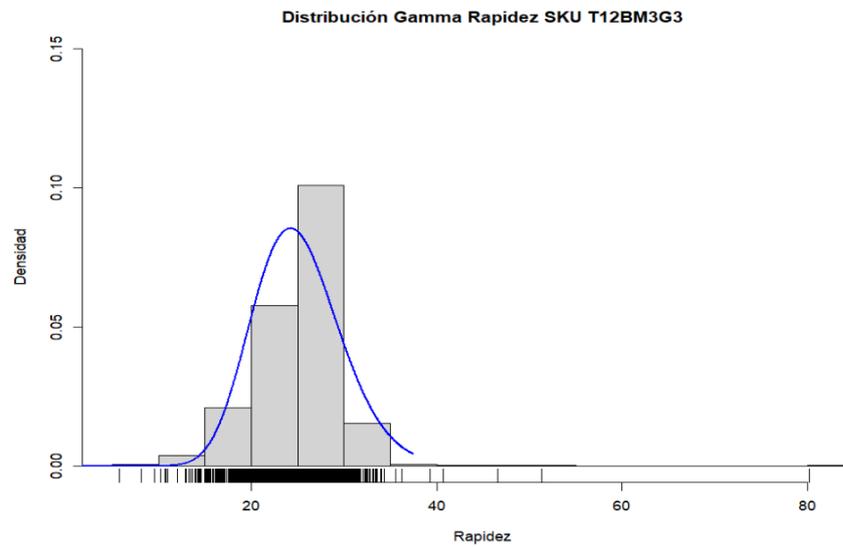


Ilustración 4: 11: Distribución Gamma Rapidez.

La frecuencia de las distribuciones de tiempo de para está dada por Ilustración

4: 12

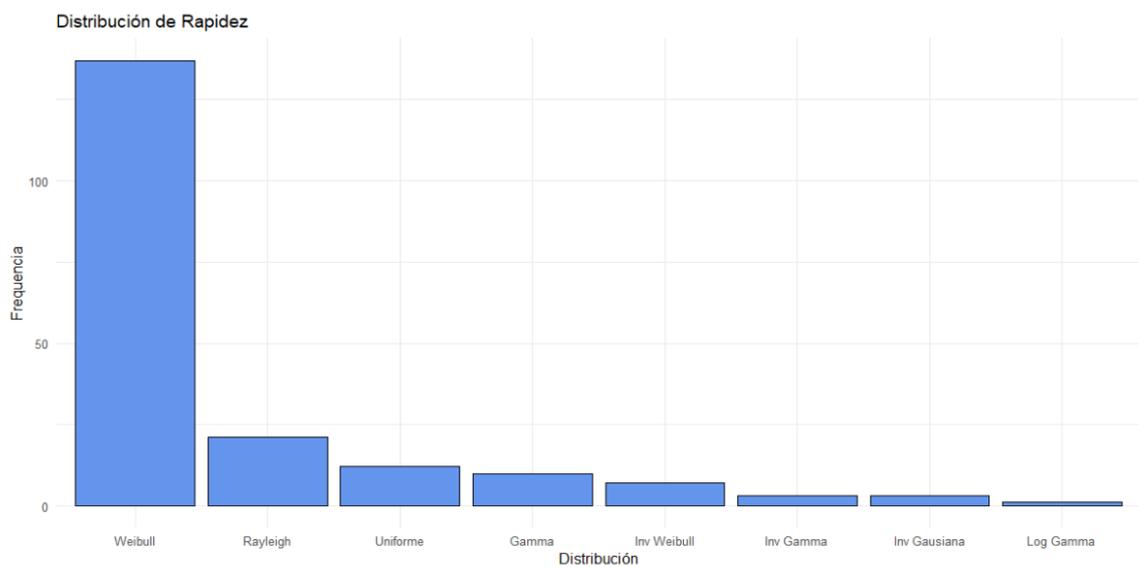


Ilustración 4: 12: Frecuencia de Distribuciones de Rapidez.

4.3. Algoritmo Generador de Secuencias.

Se generan 2372 planes de producción, a partir de los mismos se proceden a generar dos secuencias de producción, obtenidos los siguientes resultados:

4.3.1. Criterio Mínimo Tiempo de Para entre el SKU's

En Tabla 4. 1, se detallan algunas de las secuencias de producción:

#	Secuencia	Día	Turno	Tiempo Producción	Tiempo Para	Rapidez	Ton
1	T113M2GB T321R2GB T412R2GB T434R2G0 T112M2GB T113M2GW T15BQ3G0 T12BM3G3 T15AQ3G0 T434R2GB	Miércoles	Diurno	13.18	0.13	22.06	290.78
2	T120M3G3 T414R2GB T115M2GB T12BM3G3	Miércoles	Nocturno	3.85	0.05	19.01	73.10
3	T551R2GB T113R2GB T111M2GB T112M2GB T114M2GB T115M2GB T551RAGB T114R2GB T120M3G3 T12BMAG0 T12BM3G3 T541R2GB T152M2GB T214R1G0	Jueves	Diurno	11.27	0.10	20.66	232.77
:	:	:	:	:	:	:	:

237	T112N2GB						
1	T120M3G0	Lunes	Diurno	3.27	0.36	17.70	57.88
	T212R1G0						
	T124M2GB						
	T113R8GI						
	T113R8GJ						
237	T112N2GB						
2	T114M2GB	Lunes	Nocturno	5.48	0.81	22.46	122.99
	T120M3G0						
	T112N8GJ						
	T130N1GB						
	T212R1GA						
	T131N1GB						
	T138C1G0						
	T151N3G3						

Tabla 4. 1: Secuencias Criterio Mínimo Tiempo de Para.

4.3.2. Criterio Máxima Rapidez de SKU

En Tabla 4. 2, se detallan algunas de las secuencias de producción:

#	Secuencia	Día	Turno	Tiempo Producción	Tiempo Para	Rapidez	Ton
1	T113M2GB	Miércoles	Diurno	15.89	3.50	23.46	290.78
	T15AQ3G0						
	T12BM3G3						
	T113M2GW						
	T434R2GB						
	T412R2GB						
	T112M2GB						
	T434R2G0						
	T321R2GB						
	T15BQ3G0						
2	T120M3G3	Miércoles	Nocturno	4.06	0.10	18.46	73.10
	T12BM3G3						
	T414R2GB						
	T115M2GB						
3	T551R2GB	Jueves	Diurno	21.55	8.53	17.87	232.77
	T114R2GB						
	T113R2GB						
	T114M2GB						
	T112M2GB						
	T152M2GB						

	T111M2GB						
	T12BM3G3						
	T12BMAG0						
	T115M2GB						
	T551RAGB						
	T214R1G0						
	T120M3G3						
	T541R2GB						
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
237	T112N2GB	Lunes	Diurno	3.44	0.58	20.20	57.88
1	T120M3G0						
	T212R1G0						
	T113R8GI						
	T113R8GJ						
	T124M2GB						
237	T112N2GB	Lunes	Nocturno	8.77	2.91	20.99	122.99
2	T138C1G0						
	T131N1GB						
	T114M2GB						
	T112N8GJ						
	T212R1GA						
	T120M3G0						
	T151N3G3						
	T130N1GB						

Tabla 4. 2: Secuencias Criterio Máxima Rapidez de Producción.

Finalmente se procedió a comparar las dos secuencias generadas etiquetándose como óptima la secuencia que genera el mínimo tiempo de para, obteniéndose Tabla 4. 3:

Secuencia	Observaciones Óptimas	Observaciones No Óptimas	% Óptimas
Min Tiempo Para	2012	360	84.82%
Max Rapidez	360	2012	15.18%

Tabla 4. 3: Comparación Secuencias Min Tiempo Para vs Max Rapidez.

4.4. Implementación de Algoritmos de Clasificación.

4.4.1. Máquina Vector Soporte

Se procedió a generar este algoritmo de clasificación con kernels: radial y polinomial obteniéndose los siguientes resultados:

4.4.1.1. Conjunto 1: Estimación de Hiperparámetros C y γ .

Para el hiperparámetro C se probaron los valores 0.001, 0.01, 0.1, 1, 5, 10, 20; para el hiperparámetro γ se probaron los valores 0.5, 1.2, 3, 4, 5, 10, se obtuvieron

Ilustración 4: 13 Ilustración 4: 14:

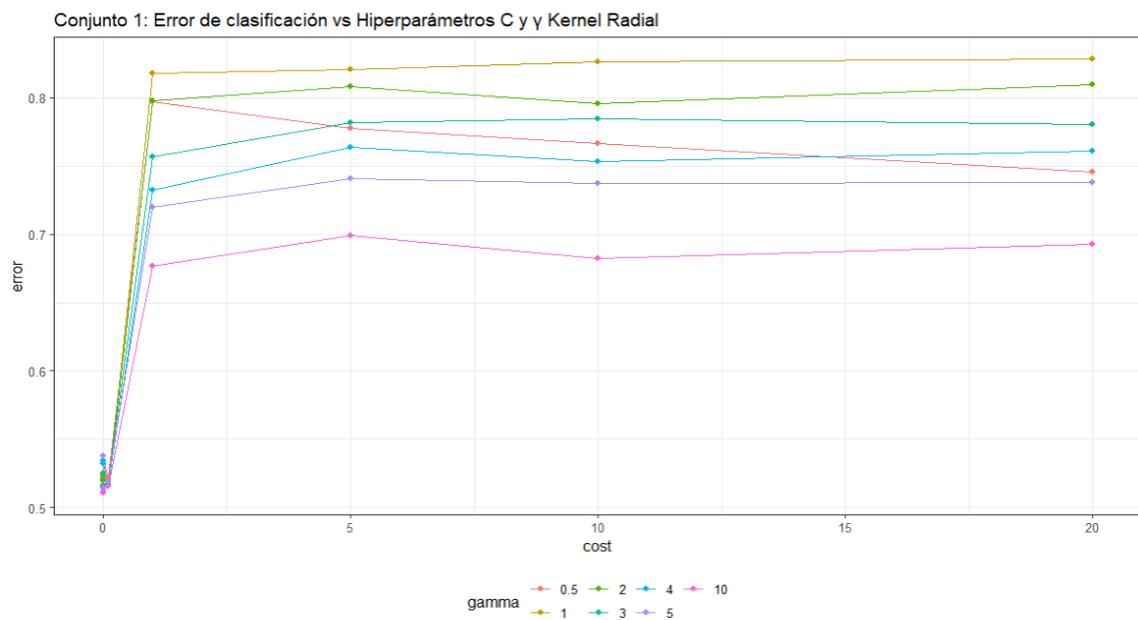


Ilustración 4: 13: Error de Clasificación vs Hiperparámetros C y γ Kernel Radial Conjunto 1.

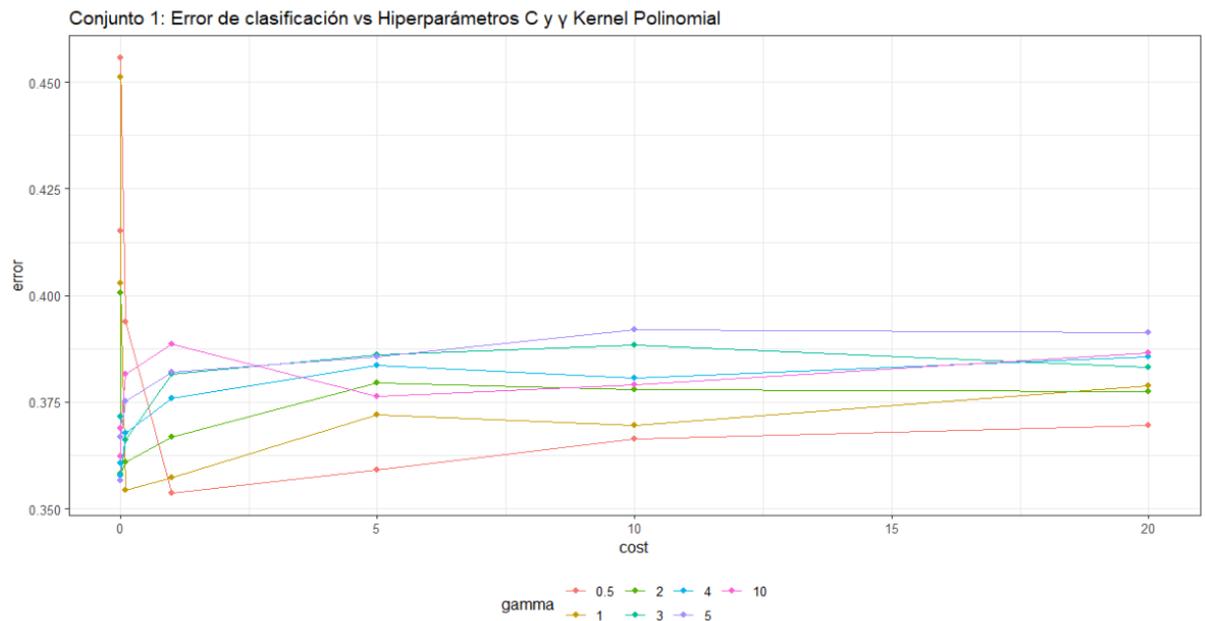


Ilustración 4: 14: Error de Clasificación vs Hiperparámetros C y γ Kernel Polinomial Conjunto 1.

En Tabla 4. 4 se detallan los mejores hiperparámetros obtenidos.

Kernel	C	γ
Radial	0.01	10
Polinomial	1	0.5

Tabla 4. 4: Hiperparámetros Máquina Vector Soporte Kernel Radial y Polinomial Conjunto 1.

Como se puede observar para el kernel radial el error crece a medida que aumenta el hiperparámetro C, luego se estabiliza para valores de C=1, independiente del valor de γ .

En cambio, para el kernel polinomial el error decrece a medida que aumenta el hiperparámetro C, alcanza su mínimo alrededor de C=0.01 y 1 y finalmente crece estabilizándose para valores de C=5.

4.4.1.2. Conjunto 1: Modelo

Una vez obtenidos los mejores hiperparámetros se procedió a ejecutar el algoritmo obteniéndose los siguientes resultados:

Grado del Kernel Polinomial: 3

Para el kernel polinomial el grado estimado se lo considera adecuado puesto que no es un grado alto.

Vectores Soporte: Los vectores de cada kernel se detallan en Tabla 4. 5

Kernel	Vectores Soporte	Vectores Totales	Proporción
Radial	3796	3796	1.0000
Polinomial	2057	3796	0.5419

Tabla 4. 5: Vectores Soporte Kernel Radial y Polinomial Conjunto 1.

Se puede apreciar que el kernel radial tiene el 100% de los vectores (observaciones) en el margen o dentro de este, este es un indicador que el modelo podría ser un muy mal predictor, respecto al kernel polinomial tiene un poco más de la mitad de los vectores dentro del margen, sigue siendo un porcentaje alto, pero se esperaría un mejor desempeño que el kernel radial.

4.4.1.3. Conjunto 1: Matriz de Confusión.

En Tabla 4. 6 Tabla 4. 7 se presenta la matriz de confusión de las predicciones para secuencias óptimas y no óptimas de los datos de entrenamiento.

		Clase Verdadera	
		Óptima	No Óptima
Clase Hipotética	Óptima	1898	1898
	No Óptima	0	0

Tabla 4. 6: Matriz de Confusión Partición Entrenamiento Kernel Radial Conjunto 1.

		Clase Verdadera	
		Óptima	No Óptima
Clase Hipotética	Óptima	1682	179
	No Óptima	216	1719

Tabla 4. 7: Matriz de Confusión Partición Entrenamiento Kernel Polinomial Conjunto 1.

Métricas: Dada la Matriz de Confusión se procede a obtener las diferentes métricas para su evaluación, las mismas se detallan en Tabla 4. 8

Métrica	Kernel Radial	Kernel Polinomial
Exactitud	0.5000	0.8959
Precisión	0.5000	0.9038
Sensibilidad	1.0000	0.8862
Especificidad	0.0000	0.9057
Prevalencia	0.5000	0.5000
Tasa de Retención	0.5000	0.4431
Retención de Prevalencia	1.0000	0.4903
Exactitud Balanceada	0.5000	0.8959
Kappa	0	0.7919

Tabla 4. 8: Métricas Partición Entrenamiento Máquina Vector Soporte Conjunto 1.

Tal como se esperaba del kernel radial está sesgado puesto que predice todas las secuencias como óptimas, en cambio el kernel polinomial tiene un mejor comportamiento.

Es importante tener en consideración que debe existir un equilibrio entre la exactitud y la precisión del modelo, en este caso ambos kernels logran este objetivo, pero a pesar de que el kernel radial logra un equilibrio exacto (exactitud 0.50 y precisión 0.50) este valor es sumamente bajo, en cambio el kernel polinomial no logra un equilibrio exacto, pero los valores de estas métricas son muy altos (exactitud 0.8959 y precisión 0.9038) algo deseable.

Adicionalmente la índice kappa nos confirma que los datos no se ajustan al kernel radial (kappa 0), pero tienen un alto ajuste al kernel polinomial (kappa 0.7919)

Inferencias: Se obtienen adicionalmente las estadísticas detalladas en Tabla

4. 9

	Kernel Radial	Kernel Polinomial
IC (95%)	(0.484, 0.516)	(0.8859, 0.9055)
Tasa Sin Información	0.50	0.50
Valor-p [Acc>NIR]	0.5065	<2 exp(-16)
Mcnemas's Test Valor-p	<2 exp(-16)	0.07009

Tabla 4. 9: Estadísticos Partición Entrenamiento Conjunto 1.

4.4.1.4. Conjunto 2: Estimación de Hiperparámetros C y γ .

Para la estimación de los hiperparámetros C y γ se realizó validación cruzada, para el hiperparámetro C se probaron los valores 0.001, 0.01, 0.1, 1, 5, 10, 20; para el hiperparámetro γ se probaron los valores 0.5, 1.2, 3, 4, 5, 10, se obtuvieron

Ilustración 4: 15 Ilustración 4: 16

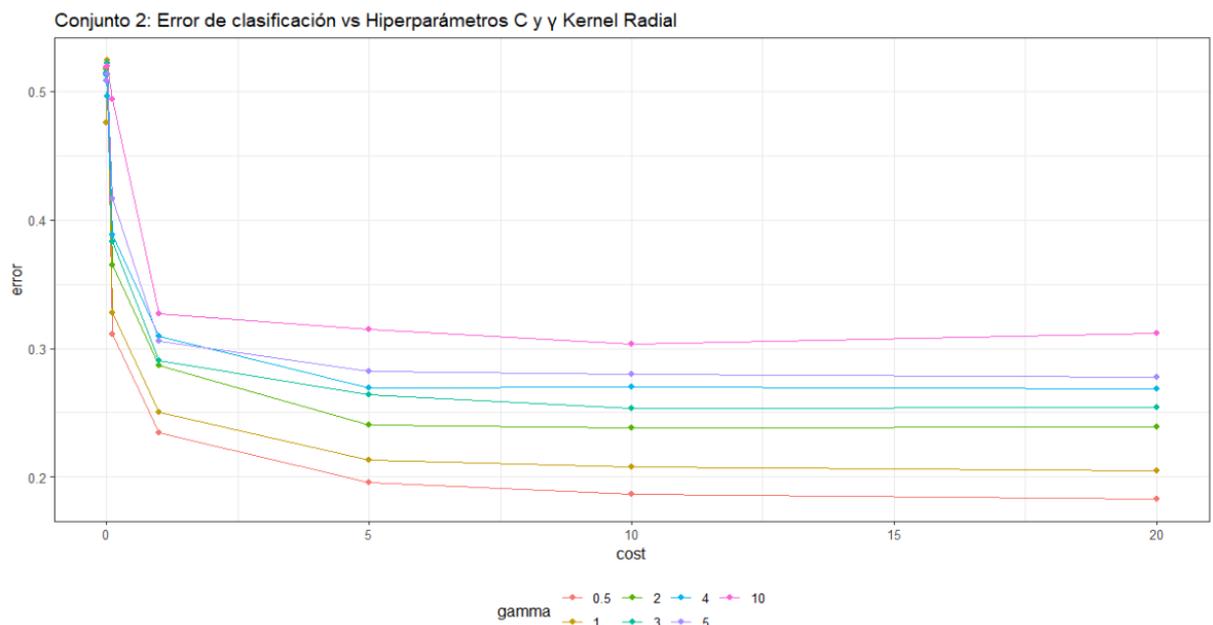


Ilustración 4: 15: Error de Clasificación vs Hiperparámetros C y γ Kernel Radial Conjunto 2.

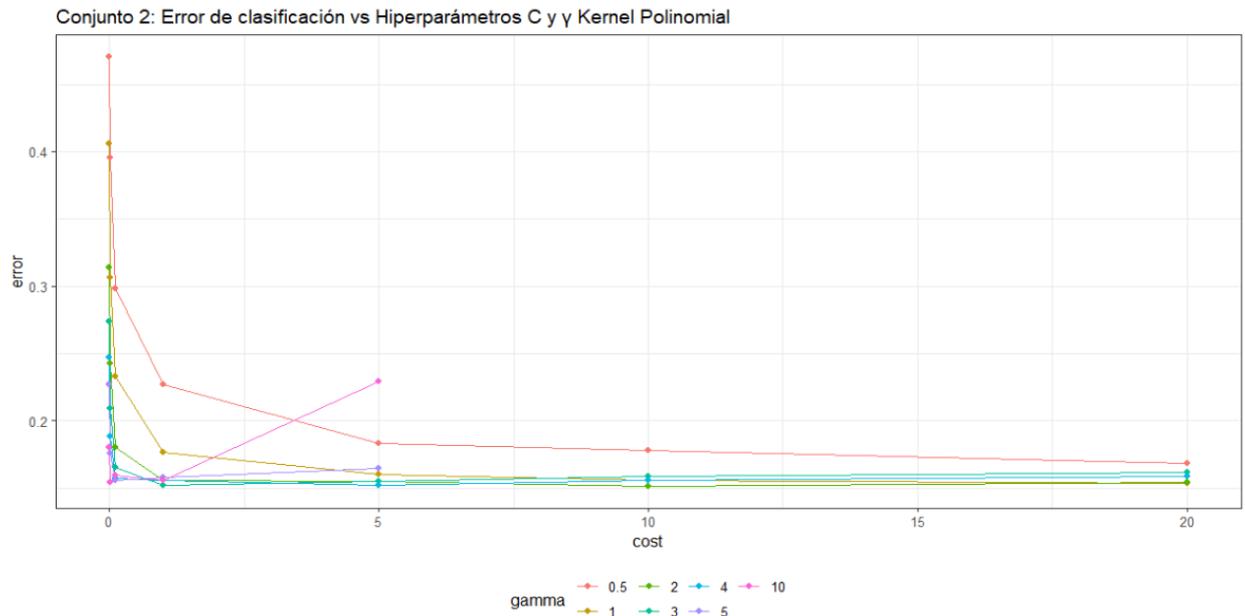


Ilustración 4. 16: Error de Clasificación vs Hiperparámetros C y γ Kernel Polinomial Conjunto 2.

En Tabla 4. 10 se detallan los mejores hiperparámetros obtenidos.

Kernel	C	γ
Radial	20	0.5
Polinomial	10	2

Tabla 4. 10: Hiperparámetros Máquina Vector Soporte Kernel Radial y Polinomial Conjunto 2.

Como se puede observar para el kernel radial que el error decrece a medida que aumenta el hiperparámetro C, luego se estabiliza para valores de C=10, independiente del valor de γ .

Con el kernel polinomial el error decrece más rápido que el kernel radial a medida que aumenta el hiperparámetro C, independiente del valor de γ y alcanza su mínimo alrededor de C=10 y 1 y finalmente crece estabilizándose para valores de C=10.

Como se puede observar en el segundo gráfico no están los valores del error para $C= 10, 20$ y $\gamma= 5, 10$, esto se debe a que a pesar de la paralelización de la máquina el algoritmo no convergió, por este motivo se procedió a ejecutar cada una de las ternas de manera individual, aquí se pudo comprobar que a medida que C y γ aumentaba el tiempo de ejecución del algoritmo crece exponencialmente puesto que para los parámetros previamente indicados el algoritmo se ejecutó durante 8 horas y no convergía. Se puede decir que el error no caerá más puesto que se puede apreciar que para valores de $\gamma = 5$ y $\gamma = 10$ el error comienza a crecer rápidamente a partir de $C=5$.

4.4.1.5. Conjunto 2: Modelo

Grado del Kernel Polinomial: 3

El grado del kernel polinomial es de 3, un grado adecuado, además es el mismo grado del kernel cuando se consideraban todas las variables indicadoras.

Vectores Soporte: Los vectores de cada kernel se detallan en Tabla 4. 11

Kernel	Vectores Soporte	Vectores Total	Proporción
Radial	1702	3796	0.4484
Polinomial	1320	3796	0.3477

Tabla 4. 11: Vectores Soporte Kernel Radial y Polinomial Conjunto 2.

Se puede apreciar que existe mayor similitud en la cantidad de vectores del margen o dentro de él para ambos kernels en comparación con los resultados que arrojó el modelo que considera todas las variables indicadoras, adicionalmente se puede observar que el kernel polinomial sigue manteniendo menos vectores del margen que el radial.

4.4.1.6. Conjunto 2: Matriz de Confusión.

En Tabla 4. 12 Tabla 4. 13 se presenta la matriz de confusión de las predicciones para secuencias óptimas y no óptimas de los datos de entrenamiento:

		Clase Verdadera	
		Óptima	No Óptima
Clase Hipotética	Óptima	1753	419
	No Óptima	145	1479

Tabla 4. 12: Matriz de Confusión Partición Entrenamiento Kernel Radial Conjunto 2.

		Clase Verdadera	
		Óptima	No Óptima
Clase Hipotética	Óptima	1774	419
	No Óptima	124	1479

Tabla 4. 13: Matriz de Confusión Partición Entrenamiento Kernel Polinomial Conjunto 2.

Métricas: Dada la Matriz de Confusión se procede a obtener las diferentes métricas para su evaluación, las mismas se detallan en Tabla 4. 14

Métrica	Kernel Radial	Kernel Polinomial
Exactitud	0.8514	0.8570
Precisión	0.8071	0.8089
Sensibilidad	0.9236	0.9347
Especificidad	0.7792	0.7792
Prevalencia	0.5000	0.5000
Tasa de Retención	0.4618	0.4673
Retención de Prevalencia	0.5722	0.5777
Exactitud Balanceada	0.8514	0.8570
Kappa	0.7028	0.7139

Tabla 4. 14: Métricas Partición Entrenamiento Máquina Vector Soporte Conjunto 2.

Se puede observar que efectivamente ambos kernels logran un equilibrio entre exactitud y precisión (0.8514, 0.8071 y 0.8570, 0.8089 respectivamente)

Adicionalmente ambos kernels el índice de kappa es muy alto y similar, mismo que indica que los datos se ajustan bien a ambos modelos.

Inferencias: Se obtienen adicionalmente las estadísticas detalladas en Tabla 4. 15

	Kernel Radial	Kernel Polinomial
IC (95%)	(0.8397, 0.8626)	(0.8454, 0.8579)
Tasa Sin Información	0.50	0.50
Valor-p [Acc>NIR]	<2.2 exp(-16)	<2.2 exp(-16)
McNemar's Test Valor-p	<2.2 exp(-16)	<2.2 exp(-16)

Tabla 4. 15: Estadísticos Partición Entrenamiento Conjunto 2.

4.4.2. Árbol de Clasificación

4.4.2.1. Conjunto 1: Modelo

Se ejecutó inicialmente el modelo con un valor de $cp=0.01$ obteniéndose que la importancia de las variables detalladas en Tabla 4. 16:

Importancia de Variable	Peso
Tiempo de Para	49
Rapidez	18
Tiempo de Producción	16
Toneladas	6
Día de Semana	3
T312R2GB	2
T312R1GA	2
T312R2GN	2
T311R1GA	2

Tabla 4. 16: Importancia de Variables Conjunto 1.

Se puede apreciar que, de las 210 variables indicadoras, solo 4 de ellas aparecen en el conjunto de variables importantes con un bajo peso, se presentan con

alto peso las variables cuantitativas, aparece la variable categórica Día de la semana, pero no la variable turno.

Luego se procedió a realizar la validación cruzada para diferentes valores de cp obteniéndose Ilustración 4: 17

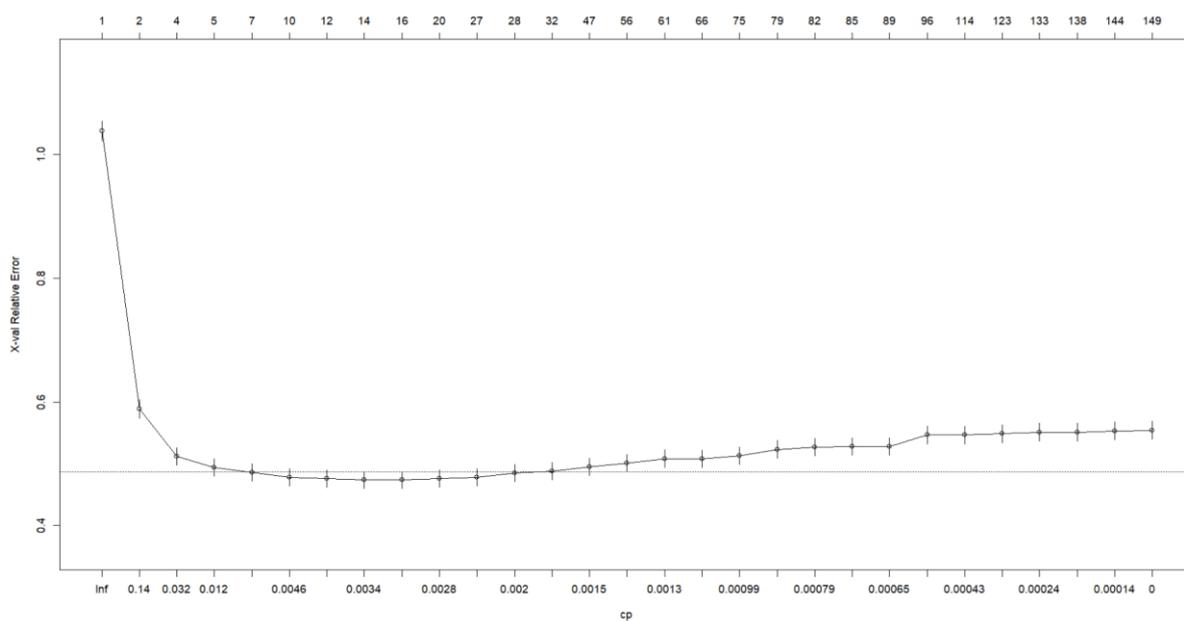


Ilustración 4: 17: Parámetro cp vs xerror.

Una vez obtenido el cp óptimo se poda el árbol inicial quedando el árbol final, el árbol final se detalla en Ilustración 4: 18

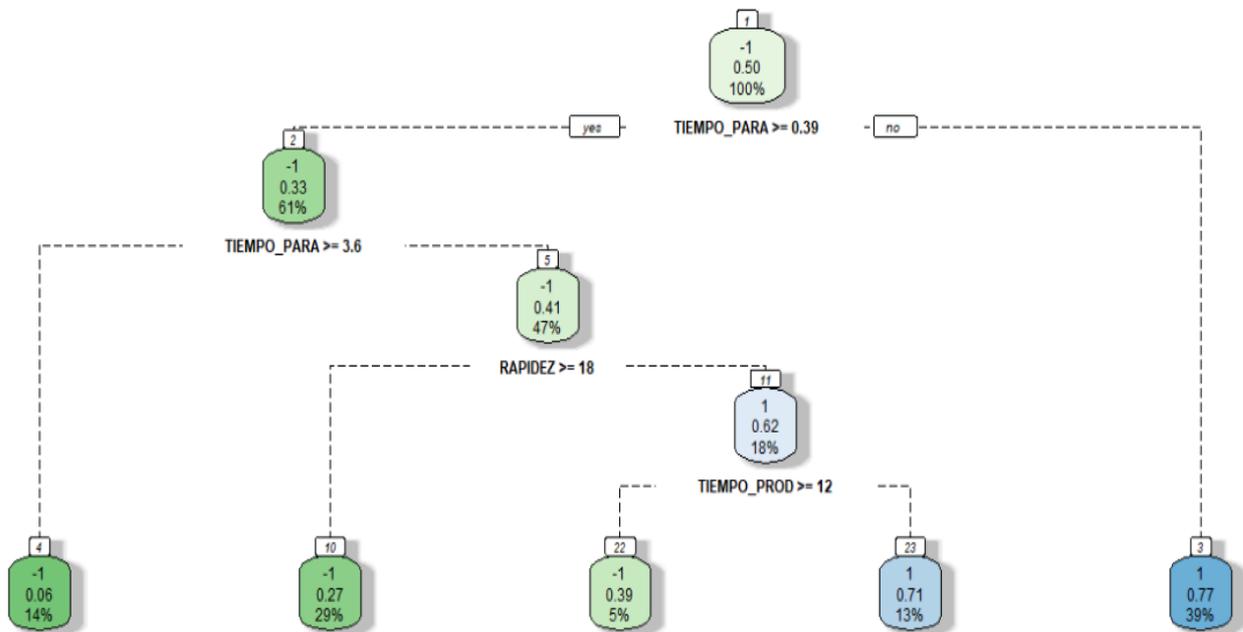


Ilustración 4: 18: Árbol de Clasificación.

Finalmente se generaron las reglas de clasificación detalladas en Tabla 4. 17

Nodo	Condición	n	Diviance
1	Raiz	3796	1898
2	Tiempo_Para >= 0.39	2316	761
4	Tiempo_Para >= 3.60	520	29
5	Tiempo_Para < 3.60	1796	732
10	Rapidez >= 18.43	1100	300
11	Rapidez < 18.43	696	264
22	Tiempo_Prod >= 12.10	199	77
23	Tiempo_Prod < 12.10	497	142
3	Tiempo_Para < 0.39	1480	343

Tabla 4. 17: Reglas de Árbol de Clasificación.

4.4.2.2. Conjunto 1: Matriz de Confusión.

En Tabla 4. 18 se presenta la matriz de confusión de las predicciones para secuencias óptimas y no óptimas de los datos de entrenamiento

		Clase Verdadera	
		Óptima	No Óptima
Clase Hipotética	Óptima	1492	485
	No Óptima	406	1413

Tabla 4. 18: Matriz de Confusión Partición Entrenamiento Árbol de Clasificación Conjunto 1.

Métricas: Dada la Matriz de Confusión se procede a obtener las diferentes métricas para su evaluación, las mismas se detallan en Tabla 4. 19

Métrica	Árbol
Exactitud	0.7653
Precisión	0.7546
Sensibilidad	0.7861
Especificidad	0.7445
Prevalencia	0.5000
Tasa de Retención	0.3930
Retención de Prevalencia	0.5208
Exactitud Balanceada	0.7653
Kappa	0.5306

Tabla 4. 19: Métricas Partición Entrenamiento Conjunto 1.

Se puede apreciar que existe un equilibrio entre exactitud y precisión, y además de buena calidad, pero el kappa nos indica que el modelo tiene concordancia moderada.

Inferencias: Se obtienen adicionalmente las estadísticas detalladas en Tabla

4. 20

IC (95%)	Tasa Sin Información	Valor-p [Acc>NIR]	McNemas's Test Valor-p
(0.7515, 0.7787)	0.50	<2.2 exp(-16)	0.008973

Tabla 4. 20: Estadísticos Partición Entrenamiento Conjunto 1.

4.4.2.3. Conjunto 2: Modelo.

Se procedió a ejecutar el algoritmo del árbol de clasificación y generó los mismos resultados que el conjunto 1, pero la importancia de las variables varió, se detalla en Tabla 4. 21

Importancia de Variable	Peso
Tiempo de Para	53
Rapidez	19
Tiempo de Producción	17
Toneladas	8
Día de Semana	3

Tabla 4. 21: Importancia de Variables Conjunto 2.

Se puede apreciar que una vez que no se cuenta con las variables el mayor peso de estas es absorbido por la variable tiempo de para.

4.4.2.4. Conjunto 2: Matriz de Confusión.

En Tabla 4. 22 se presenta la matriz de confusión de las predicciones para secuencias óptimas y no óptimas de los datos de entrenamiento.

		Real	
		Óptima	No Óptima
Predicción	Óptima	1492	485
	No Óptima	406	1413

Tabla 4. 22: Matriz de Confusión Partición Entrenamiento Árbol de Clasificación Conjunto 2.

Métricas: Dada la Matriz de Confusión se procede a obtener las diferentes métricas para su evaluación, las mismas se detallan en Tabla 4. 23

Métrica	Árbol
Exactitud	0.7653
Precisión	0.7547
Sensibilidad	0.7861
Especificidad	0.7445
Prevalencia	0.5000
Tasa de Retención	0.3930
Retención de Prevalencia	0.5208
Exactitud Balanceada	0.7653
Kappa	0.5306

Tabla 4. 23: Métricas Partición Entrenamiento Conjunto 2

Al igual que en el conjunto 1, se logra un equilibrio entre exactitud y precisión, y de la misma manera el kappa indica una moderada concordancia.

Inferencias: Se obtienen adicionalmente las estadísticas detalladas en Tabla 4. 24

IC (95%)	Tasa Sin Información	Valor-p [Acc>NIR]	Mcnemas's Test Valor-p
(0.7515, 0.7787)	0.50	<2.2 exp(-16)	0.008973

Tabla 4. 24: Estadísticos Partición Entrenamiento Conjunto 2.

4.4.3. Comparación de Algoritmos

4.4.3.1. Métricas

Para comparar los algoritmos se tomaron en cuentas las métricas de exactitud, precisión y kappa obteniéndose Ilustración 4: 24

Conjunto	Algoritmo	Exactitud	Precisión	Kappa
Conjunto 1	Kernel Radial	0.5000	0.5000	0.0000
Conjunto 1	Kernel Polinomial	0.8959	0.9038	0.7959
Conjunto 1	Árbol Clasificación	0.7653	0.7546	0.5306
Conjunto 2	Kernel Radial	0.8514	0.8071	0.7028
Conjunto 2	Kernel Polinomial	0.8570	0.8089	0.7139
Conjunto 2	Árbol Clasificación	0.7653	0.7546	0.5306

Tabla 4. 25: Métricas Partición Entrenamiento por Algoritmo y Conjunto

De manera gráfica se obtiene Ilustración 4: 19 Ilustración 4: 20 Ilustración 4:

21:

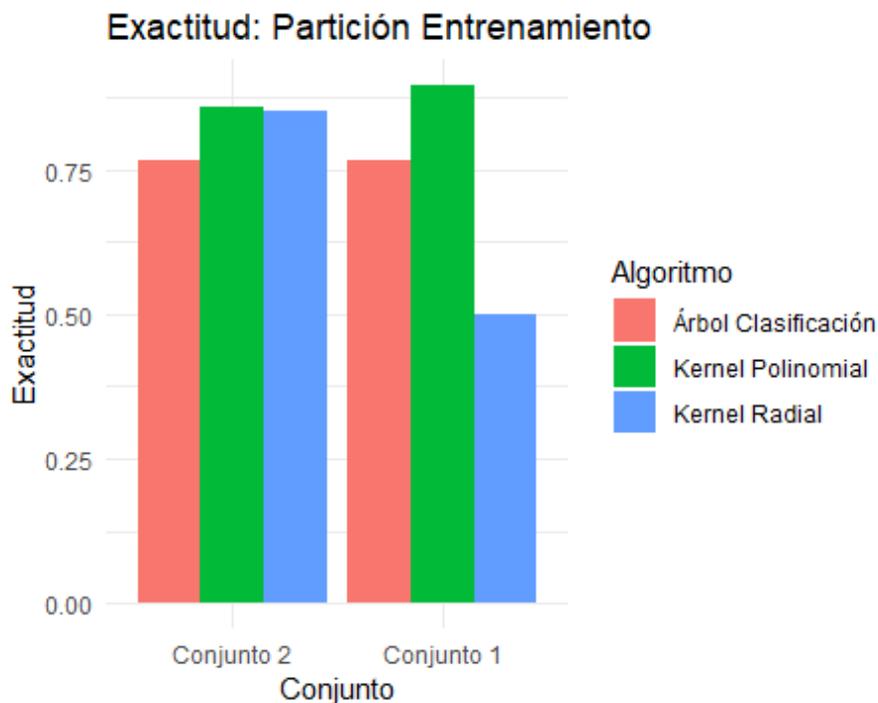


Ilustración 4: 19: Exactitud Partición Entrenamiento por Algoritmo y Conjunto.

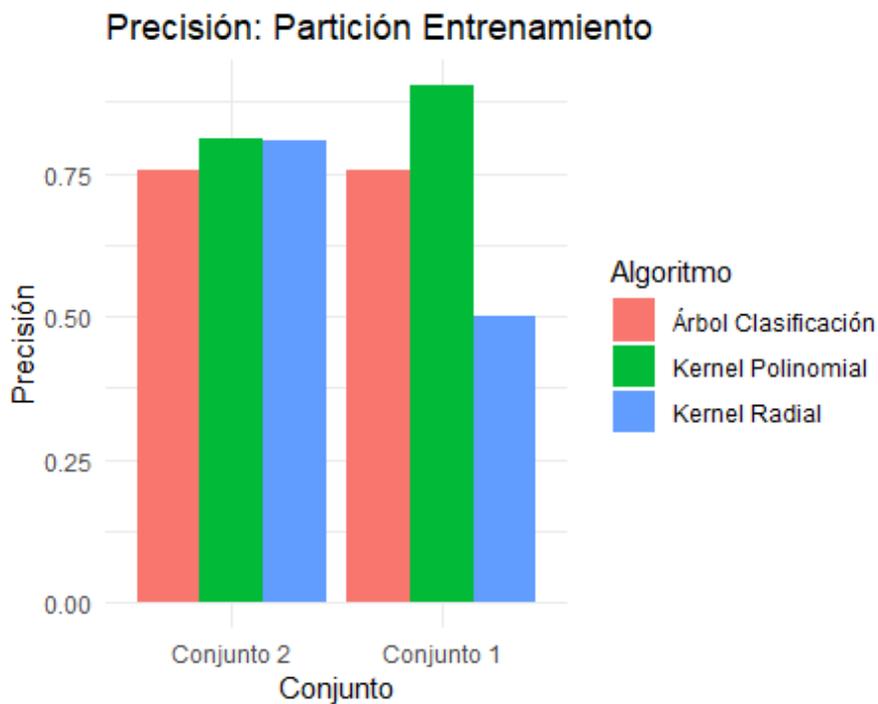


Ilustración 4: 20: Precisión Partición Entrenamiento por Algoritmo y Conjunto.

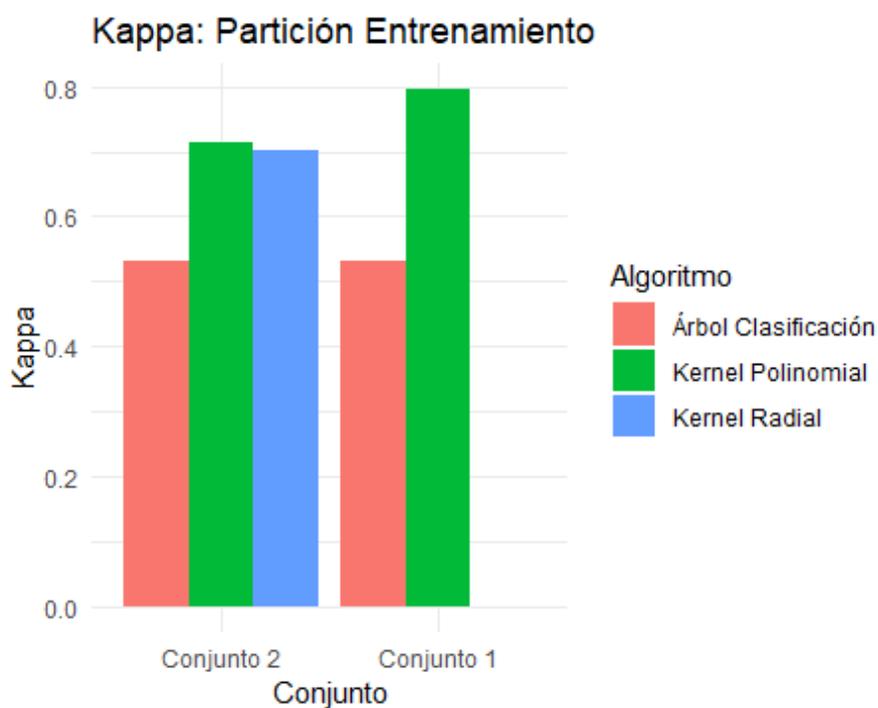


Ilustración 4: 21: Kappa Partición Entrenamiento por Algoritmo y Conjunto.

Se puede apreciar que el algoritmo con kernel radial para el conjunto que contiene todas las variables (Conjunto 1) tiene un bajo valor para todas las métricas, en cambio el kernel polinomial tiene un muy buen desempeño en este conjunto de datos, el árbol de decisión tiene un desempeño superior al kernel radial pero inferior al polinomial.

Para el conjunto que excluye las variables indicadoras (Conjunto 2) el algoritmo con kernel radial tiene una mejora sustancial en su desempeño para todas las métricas, en cambio el algoritmo con kernel polinomial baja su desempeño, finalmente el árbol de decisión mantiene el mismo rendimiento para ambos conjuntos.

Según el estadístico Kappa se podría decir que para el kernel radial aplicado al conjunto 1 la clasificación se realizó al azar puesto que su valor es 0, este estadístico indica buena concordancia cuando se aplica este modelo al conjunto 2.

Para el kernel polinomial para ambos conjuntos se dice que el estadístico kappa indica buena concordancia para ambos conjuntos, aunque decae una décima en el conjunto 2.

Para el árbol de clasificación su concordancia es moderada para ambos conjuntos de datos.

4.4.3.2. Tiempos de ejecución

También se midió el tiempo de ejecución en la calibración de los modelos, teniendo en consideración que para el algoritmo Máquina Vector Soporte se evaluó el tiempo de ejecución en evaluar los 49 pares de valores de C y γ mientras que para

el algoritmo árbol de clasificación se midió el tiempo de podado del árbol inicial, obteniéndose Tabla 4. 25

Algoritmo	Conjunto	Tiempo de Ejecución (Horas)
MVS Kernel Radial	Conjunto 1	1.14791400
MVS Kernel Polinomial	Conjunto 1	0.70853600
Árbol de Clasificación	Conjunto 1	0.00076816
MVS Kernel Radial	Conjunto 2	0.46868100
MVS Kernel Polinomial	Conjunto 2	∞
Árbol de Clasificación	Conjunto 2	0.00034532

Tabla 4. 26: Tiempo de ejecución Partición Entrenamiento por Algoritmo y Conjunto.

De manera gráfica se obtiene Ilustración 4: 22:

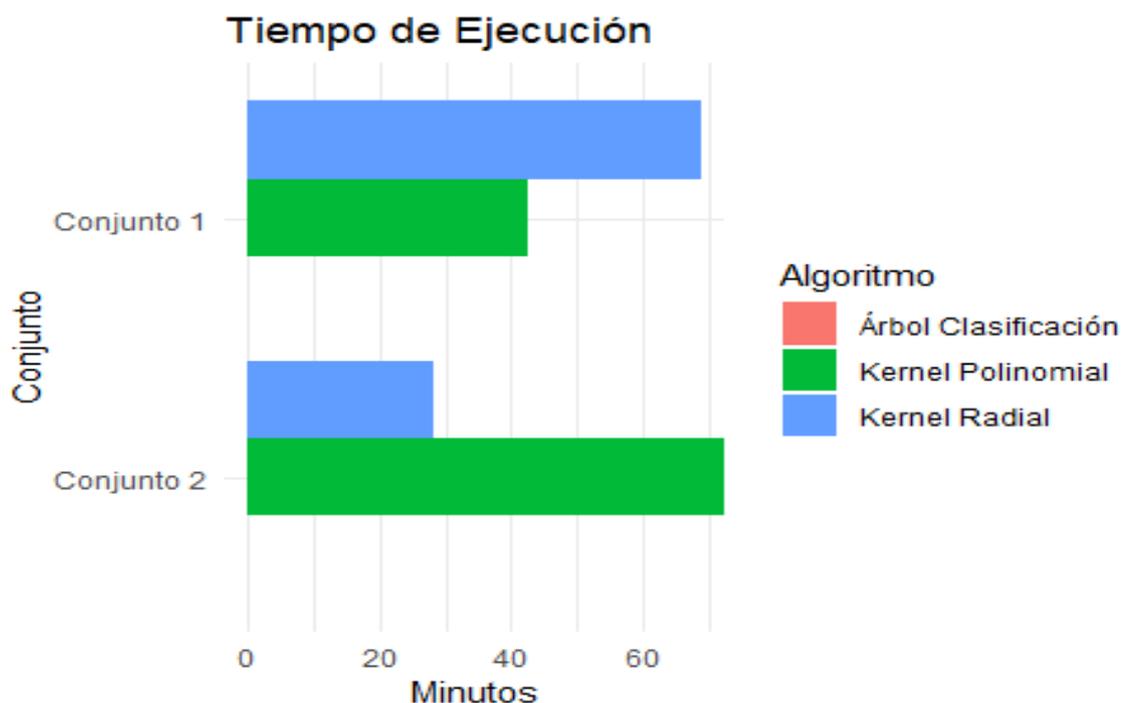


Ilustración 4: 22: Tiempo de ejecución Partición Entrenamiento por Algoritmo y Conjunto.

Se puede apreciar que el kernel radial para el conjunto que contiene las variables indicadoras tiene un alto tiempo de ejecución, en cambio el mismo se reduce

casi a la mitad cuando se trabaja sin variables indicadoras, ahora si se recuerda las métricas de exactitud, precisión y kappa estas mejoraban sustancialmente cuando se las media sobre el conjunto que no contenía las variables indicadoras, es decir mayor tiempo de ejecución no implica un mejor modelo para el kernel radial.

Para el modelo con kernel polinomial el tiempo de ejecución crece de manera exponencial, de hecho, de los 49 pares de hiperparámetros C y γ solo se pudo estimar 45 pares, los últimos 4 pares corresponden a los valores más altos C y γ y los mismos no pudieron ser estimados puesto que transcurrieron más de 15 horas y no convergieron ninguno de ellos, este caso se contradice la idea intuitiva que a menor número de variables se incurriría menor número de ejecución para el kernel polinomial.

Respecto al árbol de clasificación el tiempo de ejecución es prácticamente el mismo y cercado al segundo, si bien es cierto el computo de los algoritmos MVS y árbol de clasificación no es mismo, si se analiza el conjunto de entrenamiento, se puede decir que para una exactitud alrededor de 0.75 un tiempo de ejecución inferior a 2 segundos es muy bueno si se compara a una exactitud alrededor del 0.87 y 0.67 para el kernel polinomial y radial respectivamente, donde el tiempo de ejecución del kernel polinomial no es viable ni existe certeza de convergencia y el tiempo de ejecución del kernel radial es de 0.8 horas aproximadamente.

4.5. Evaluación de Algoritmos de Clasificación.

Se procede a implementar los algoritmos en los datos de entrenamiento, obteniéndose los siguientes resultados:

4.5.1. Máquina Vector Soporte

4.5.1.1. Conjunto 1: Matriz de Confusión.

En Tabla 4. 27 Tabla 4. 28 se presenta la matriz de confusión de las predicciones para secuencias óptimas y no óptimas de los datos de prueba:

	Clase Verdadera	
	Óptima	No Óptima
Clase	474	474
Hipotética	0	0

Tabla 4. 27: Matriz de Confusión Partición Prueba Kernel Radial Conjunto 1.

	Clase Verdadera	
	Óptima	No Óptima
Clase	Óptima 427	45
Hipotética	No Óptima 47	429

Tabla 4. 28: Matriz de Confusión Partición Prueba Kernel Polinomial Conjunto 1.

Métricas: Dada la Matriz de Confusión se procede a obtener las diferentes métricas para su evaluación, las mismas se detallan en Tabla 4. 29

Métrica	Kernel Radial	Kernel Polinomial
Exactitud	0.5000	0.903
Precisión	0.5000	0.9047
Sensibilidad	1.0000	0.9008
Especificidad	0.0000	0.9051
Prevalencia	0.5000	0.5000
Tasa de Retención	0.5000	0.4504
Retención de Prevalencia	1.0000	0.4979
Exactitud Balanceada	0.5000	0.9030
Kappa	0	0.8059

Tabla 4. 29: Métricas Partición Prueba Máquina Vector Soporte Conjunto 1.

Inferencias: Se obtienen adicionalmente las estadísticas detalladas en Tabla

4. 30

	Kernel Radial	Kernel Polinomial
IC (95%)	(0.4677, 0.5323)	(0.882, 0.821)
Tasa Sin Información	0.50	0.50
Valor-p [Acc>NIR]	0.513	<2 exp(-16)
Mcnemas's Test Valor-p	<2 exp(-16)	0.917

Tabla 4. 30: Estadísticos Partición Prueba Máquina Vector Soporte Conjunto 1.

4.5.1.2. Conjunto 2: Matriz de Confusión.

En Tabla 4. 31 Tabla 4. 32 se presenta la matriz de confusión de las predicciones para secuencias óptimas y no óptimas de los datos de prueba

		Clase Verdadera	
		Óptima	No Óptima
Clase Hipotética	Óptima	448	101
	No Óptima	26	373

Tabla 4. 31: Matriz de Confusión Partición Prueba Kernel Radial Conjunto 2.

		Clase Verdadera	
		Óptima	No Óptima
Clase Hipotética	Óptima	446	113
	No Óptima	28	361

Tabla 4. 32: Matriz de Confusión Partición Prueba Kernel Polinomial Conjunto 2.

Métricas: Dada la Matriz de Confusión se procede a obtener las diferentes métricas para su evaluación, las mismas se detallan en Tabla 4. 33

Métrica	Kernel Radial	Kernel Polinomial
Exactitud	0.8660	0.8513
Precisión	0.8160	0.7979
Sensibilidad	0.9451	0.9409
Especificidad	0.7869	0.7616
Prevalencia	0.5000	0.5000
Tasa de Retención	0.4726	0.4705
Retención de Prevalencia	0.5791	0.5897
Exactitud Balanceada	0.8660	0.8513
Kappa	0.7321	0.7025

Tabla 4. 33: Métricas Partición Prueba Máquina Vector Soporte Conjunto 2.

Inferencias: Se obtienen adicionalmente las estadísticas, las mismas se detallan en Tabla 4. 34

	Kernel Radial	Kernel Polinomial
IC (95%)	(0.8427, 0.8871)	(0.827, 0.8733)
Tasa Sin Información	0.50	0.50
Valor-p [Acc>NIR]	<2.200 exp(-16)	<2.200 exp(-16)
McNemans's Test Valor-p	<5.153 exp(-11)	1.504 exp(-16)

Tabla 4. 34: Estadísticos Partición Prueba Máquina Vector Soporte Conjunto 1.

4.5.2. Árbol de Clasificación

4.5.2.1. Conjunto 1: Matriz de Confusión:

En Tabla 4. 35 se presenta la matriz de confusión de las predicciones para secuencias óptimas y no óptimas de los datos de prueba.

		Clase Verdadera	
		Óptima	No Óptima
Clase Hipotética	Óptima	368	127
	No Óptima	106	347

Tabla 4. 35: Matriz de Confusión Partición Prueba Árbol de Clasificación Conjunto 1

Métricas: Dada la Matriz de Confusión se procede a obtener las diferentes métricas para su evaluación, las mismas se detallan en Tabla 4. 36

Métrica	Árbol
Exactitud	0.7542
Precisión	0.7434
Sensibilidad	0.7764
Especificidad	0.7321
Prevalencia	0.5000
Tasa de Retención	0.3882
Retención de Prevalencia	0.5222
Exactitud Balanceada	0.7542
Kappa	0.5084

Tabla 4. 36: Matriz de Confusión Partición Prueba Árbol de Clasificación Conjunto 1.

Inferencias: Se obtienen adicionalmente las estadísticas detalladas en Tabla 4. 37

IC (95%)	Tasa Sin Información	Valor-p [Acc>NIR]	Mcnemas's Test Valor-p
(0.7255, 0.7813)	0.50	<2 exp(-16)	0.1901

Tabla 4. 37: Estadísticos Partición Prueba Árbol de Clasificación Conjunto 1.

4.5.2.2. Conjunto 2: Matriz de Confusión.

En Tabla 4. 38 se presenta la matriz de confusión de las predicciones para secuencias óptimas y no óptimas de los datos de prueba:

		Clase Verdadera	
		Óptima	No Óptima
Clase Hipotética	Óptima	368	127
	No Óptima	106	347

Tabla 4. 38: Matriz de Confusión Partición Prueba Árbol de Clasificación Conjunto 2.

Métricas: Dada la Matriz de Confusión se procede a obtener las diferentes métricas para su evaluación, las mismas se detallan en Tabla 4. 39

Métrica	Árbol
Exactitud	0.7542
Precisión	0.7434
Sensibilidad	0.7764
Especificidad	0.7321
Prevalencia	0.5000
Tasa de Retención	0.3882
Retención de Prevalencia	0.5222
Exactitud Balanceada	0.7542
Kappa	0.5084

Tabla 4. 39: Métricas Partición Prueba Árbol de Clasificación Conjunto 2.

Inferencias: Se obtienen adicionalmente las estadísticas detalladas en Tabla

4. 40

IC (95%)	Tasa Sin Información	Valor-p [Acc>NIR]	McNemas's Test Valor-p
(0.7255, 0.7813)	0.50	<2 exp(-16)	0.1901

Tabla 4. 40: Estadísticos Partición Prueba Árbol de Clasificación Conjunto 2.

4.5.3. Comparación de Algoritmos.

Para comparar el desempeño de los algoritmos se evaluó su predicción para el conjunto de prueba y se consideraron las métricas de exactitud, precisión y kappa obteniéndose los resultados detallados en Tabla 4. 41

Conjunto	Algoritmo	Exactitud	Precisión	Kappa
Conjunto 1	Kernel Radial	0.5000	0.5000	0.0000
Conjunto 1	Kernel Polinomial	0.9030	0.9047	0.8059
Conjunto 1	Árbol Clasificación	0.7542	0.7434	0.5084
Conjunto 2	Kernel Radial	0.8660	0.8160	0.7321
Conjunto 2	Kernel Polinomial	0.8513	0.7979	0.7025
Conjunto 2	Árbol Clasificación	0.7542	0.7434	0.5084

Tabla 4. 41: Métricas Partición Prueba por Algoritmo y Conjunto.

De manera gráfica se obtiene Ilustración 4: 23 Ilustración 4: 24 Ilustración 4:

25:

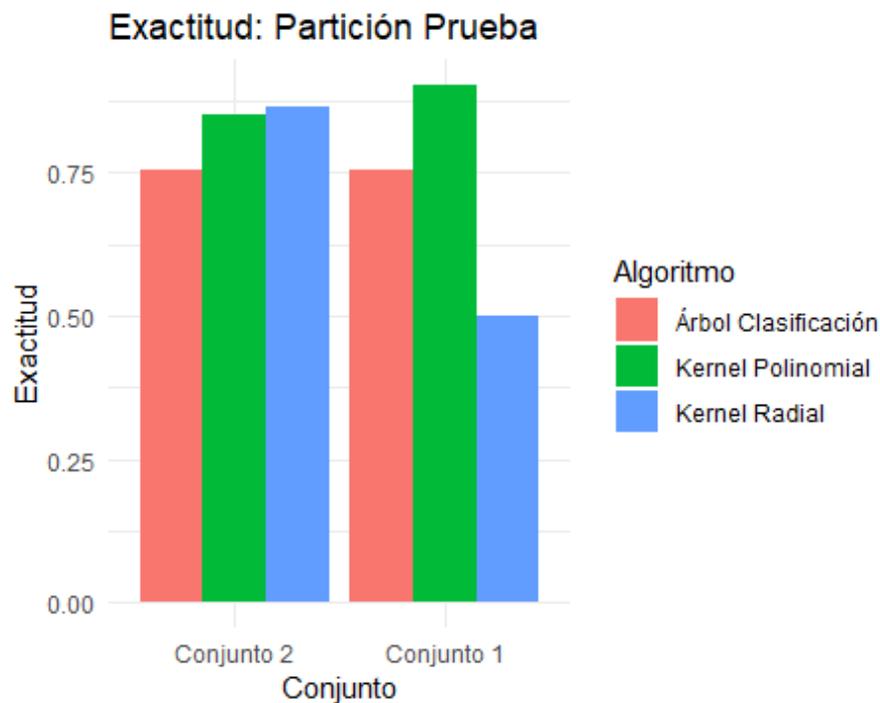


Ilustración 4: 23: Exactitud Partición Prueba por Algoritmo y Conjunto.

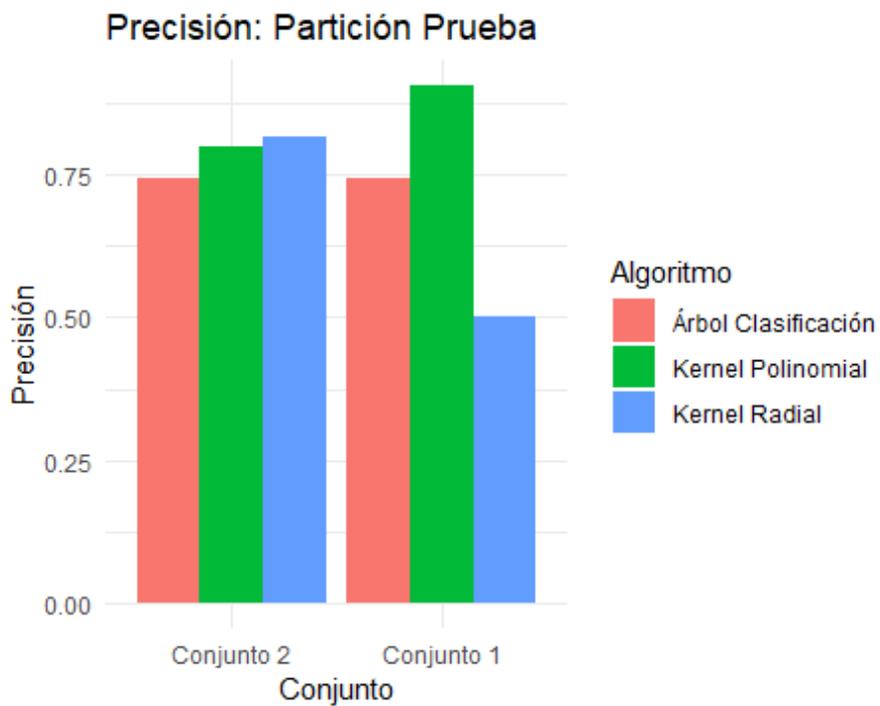


Ilustración 4: 24: Precisión Partición Prueba por Algoritmo y Conjunto.

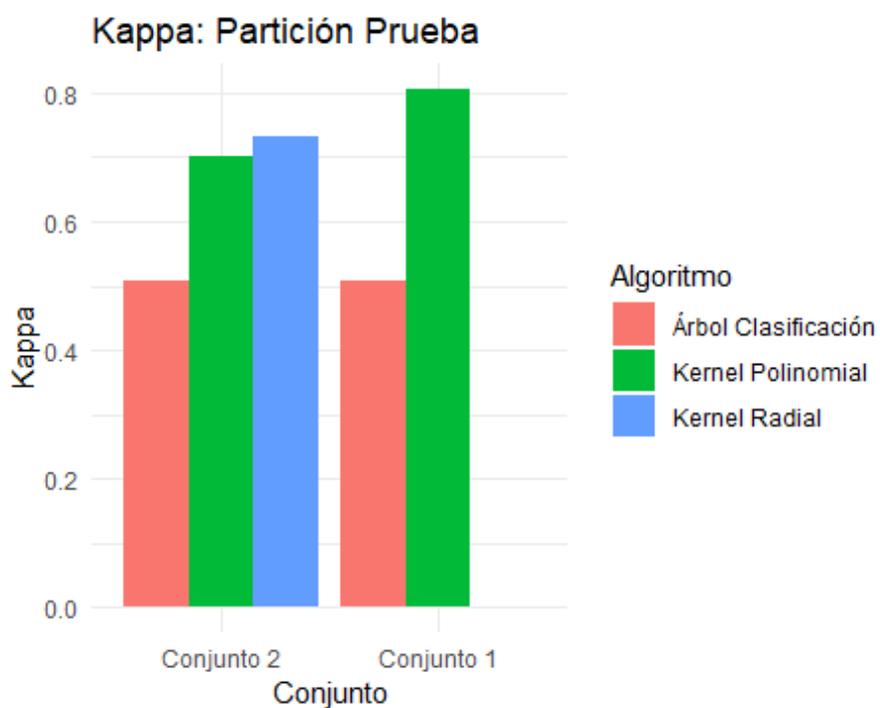


Ilustración 4: 25: Kappa Partición Prueba por Algoritmo y Conjunto.

Se puede apreciar que el comportamiento de los algoritmos en los datos de prueba es muy similar al comportamiento de los datos de prueba tanto ambos conjuntos, pero para el conjunto sin variables indicadoras existe una variación en los algoritmos de kernel radial y polinomial, puesto que en el conjunto de entrenamiento el kernel polinomial tenía un rendimiento ligeramente superior al kernel radial, pero en el conjunto de prueba el rendimiento del kernel radial supera en poco al rendimiento del kernel polinomial. El rendimiento del árbol de decisión no tiene variación significativa en su desempeño si se compara con el desempeño en el conjunto de entrenamiento.

Se aprecia que se mantiene el mismo estadístico $\kappa=0$ para el modelo radial cuando se aplica en el conjunto 1, para el kernel polinomial el estadístico κ indica que se tiene una muy buena concordancia, además existe una mejora respecto al κ de los datos de entrenamiento.

Para el conjunto 2 el estadístico κ del kernel radial mejora ligeramente respecto al de los datos de entrenamiento y se cataloga como muy buena concordancia. Respecto al kernel polinomial, el estadístico κ baja ligeramente respecto a los datos de entrenamiento, pero se mantiene como buena concordancia.

Para el árbol de clasificación el estadístico κ baja un poco respecto a los datos de entrenamiento, pero se sigue manteniendo una concordancia moderada.

CAPÍTULO 5

5. Conclusiones y Recomendaciones

Una vez analizados los algoritmos se pueden analizar un análisis de las fortalezas, oportunidades, debilidades y amenazas (FODA) de estos, mismo que se detalla a continuación:

5.1. Fortalezas:

El algoritmo Máquina Vector Soporte de kernel polinomial tiene un excelente rendimiento con pocas o muchas variables categóricas, los de kernel radial en cambio obtienen así mismo un excelente rendimiento cuando se trabaja con más variables cuantitativas que categóricas.

Los algoritmos Máquina Vector Soporte son altamente fiables en su rendimiento para predecir la categoría de nuevas observaciones.

El algoritmo Árbol de Clasificación son poco sensible en su rendimiento respecto al número de variables categóricas.

El tiempo computacional de ejecución del árbol de clasificación es muy bajo y esto lo convierte en una excelente herramienta a la hora de tomar decisiones operativas como la secuenciación de producción.

Las reglas del árbol de clasificación permiten una fácil interpretación por parte del usuario y no requiere un conocimiento estadístico especializado.

5.2. Oportunidades

La secuencia de las órdenes de producción fue modelada mediante variables indicadoras, es de interés analizar el desempeño de los algoritmos previamente descritos, pero modelizando las secuencias de las órdenes de producción mediante el análisis funcional de datos.

De poderse recolectar más secuencias de producción es posible definir por cada SKU una función distribución de probabilidad bivariada de kilogramos y tiempo de producción con la finalidad de tener una mejor aproximación al tiempo de producción de la secuencia.

Los algoritmos fueron ejecutados localmente, en trabajos futuros podría ejecutarse los mismos en plataformas como Google Colab o RStudio Cloud con la finalidad de reducir los tiempos de ejecución específicamente para los algoritmos de Máquina Vector Soporte ya que son los de mejor rendimiento, pero en tiempo computacional son poco viables.

En trabajos futuros se podría evaluar la implementación de Bosques Aleatorios para mejorar el desempeño de los árboles de clasificación.

Con la finalidad de mejorar los resultados del presente proyecto se ve como una oportunidad la implementación de la metodología usada a la empresa automatizando la captura de los datos que construyen las funciones de distribuciones de probabilidad de esta manera se obtendría datos en tiempo real del comportamiento del sistema.

5.3. Debilidades

El algoritmo Máquina Vector Soporte con kernel radial se ve afectado fuertemente en su desempeño con un alto volumen de variables indicadoras y clasifica todas las observaciones en una única clase.

Debido al tiempo computacional los algoritmos de Máquina Vector Soporte tienen que ser implementados usando ejecución en paralelo, ejecutados en la nube u otra estrategia que lo vuelva viable para tomar decisiones operativas sobre la secuenciación de la producción.

El algoritmo de Árbol de Clasificación tiene un comportamiento muy estable en su desempeño, pero este se podría considerar como moderado en comparación a los demás algoritmos.

5.4. Amenazas

El algoritmo de Máquina Vector Soporte para kernel polinomial puede no converger en un tiempo adecuado.

De no podarse correctamente el árbol de clasificación podría incurrirse en reglas de muy compleja interpretación por parte del usuario.

Ambos tipos de algoritmos están sujetos al posible sobreajuste a los datos de entrenamiento, pero para los datos de entrenamiento probados se corroboró que no existió sobreajuste.

6. Referencias

Referencias

- Ahn, G., Park, M., Park, Y.-J., & Hur, S. (2019). Interactive Q-Learning Approach for Pick-and-Place Optimization of the Die Attach Process in the Semiconductor Industry. *Hiundawi*.
- Alltech. (2020). *Encuesta Global Anual sobre Alimento Balanceado 2020*. LEXINGTON: Alltech.
- ALLTECH. (2020). *www.alltech.com*. Obtenido de <https://www.alltech.com/es-mx/press-release/la-encuesta-global-sobre-alimento-balanceado-de-alltech-revela-por-primera-vez-en>
- Hastie, T., Tibshirani, R., & Friedman, J. (2018). *The Elements of Statistical Learning*. Springer.
- James, G., Witten, D., Hastie, T., & Tibshirami, R. (2013). *An Introduction to Statistical Learning with Applications in R*. Springer.

Lieber, D., Stolpe, M., Konrad, B., Deuse, J., & Morik, K. (2013). Quality Prediction in Interlinked Manufacturing Processes based on Supervised & Unsupervised Machine Learning. *Elsevier*.

PROEQUADOR. (2020). www.proecuador.gob.ec. Obtenido de <https://www.proecuador.gob.ec/alimentos-para-animales/>

Reboiro-Jato, M., Glez-Dopazo, J., D, G., Rosalía, L., Juan Fco, G., Pavón, R., . . . Fdez-Riverola, F. (2011). Using inductive learning to assess compound feed production in cooperative poultry farms. *Elsevier*.

Revista Vistazo. (31 de Julio de 2020). www.vistazo.com. Obtenido de <https://www.vistazo.com/actualidad/nacional/ascenso-de-la-industria-de-alimentos-para-mascotas-FXVI192601>

Sutton, R., & Barto, A. (2018). *Reinforcement Learning An Introduction*. MIT Press.

Wickham, H., & Golemund, G. (2017). *R for Data Science*. O'Reilly Media.

Wuest, T., Weimer, D., Irgens, C., & Klaus-Dieter, T. (2016). Machine Learning in Manufacturing: Advantages, Challenges, and Applications. Production & Manufacturing Research. *Scientific Research*.

7. Apéndice y Anexos

7.1. Apéndice A: Tabla de Medidas de Tendencia Central Análisis Univariante Cuantitativo.

Presentación ¡Error! No se encuentra el origen de la referencia. ¡Error! No se encuentra el origen de la referencia. ¡Error! No se encuentra el origen de la referencia.

	DIF_HORA	KG	RAP_TON_HORA	TIME_PRE
Media	0.99	20,955.79	22.27	0.23
Mediana	0.66	14,430.96	22.58	0.15
Moda	0.34	6,952.49	22.79	0.09
Varianza	0.95	410,007,472.66	30.64	0.07
Desv Típica	0.98	20,248.64	5.54	0.26
Mínimo	0.03	1,269.04	1.06	0.03
Máximo	8.55	137,707.10	65.29	6.34

Tabla 7. 1: Medidas de Tendencia Central y Medias de Dispersión: Presentación Ensacado.

	DIF_HORA	KG	RAP_TON_HORA	TIME_PRE
Media	0.59	14,112.97	25.16	0.22
Mediana	0.49	12,586.94	25.93	0.13

Moda	0.45	12,575.99	26.84	0.09
Varianza	0.09	36,979,470.32	20.72	0.07
Desv Típica	0.30	6,081.07	4.55	0.27
Mínimo	0.02	1,594.76	1.40	0.04
Máximo	4.90	79,207.62	83.82	7.02

Tabla 7. 2: Medidas de Tendencia Central y Medias de Dispersión: Presentación Granel.

	DIF_HORA	KG	RAP_TON_HORA	TIME_PRE
Media	0.42	8,283.37	22.40	0.26
Mediana	0.35	7,192.38	23.37	0.17
Moda	0.18	3,936.33	25.22	0.10
Varianza	0.17	34,923,076.17	46.04	0.11
Desv Típica	0.41	5,909.57	6.79	0.34
Mínimo	0.03	1,355.18	0.59	0.05
Máximo	6.80	39,683.26	53.57	5.91

Tabla 7. 3: Medidas de Tendencia Central y Medias de Dispersión: Presentación Reproceso.

Tipo Proceso ¡Error! No se encuentra el origen de la referencia. ¡Error! No se encuentra el origen de la referencia. ¡Error! No se encuentra el origen de la referencia.

	DIF_HORA	KG	RAP_TON_HORA	TIME_PRE
Media	0.54	12,599.01	24.31	0.22
Mediana	0.47	12,577.61	25.13	0.14
Moda	0.44	11,766.00	26.22	0.09
Varianza	0.09	40,936,891.33	25.87	0.06
Desv Típica	0.31	6,398.19	5.09	0.25
Mínimo	0.02	1,591.02	1.06	0.04
Máximo	4.22	64,054.22	83.82	7.02

Tabla 7. 4: Medidas de Tendencia Central y Medias de Dispersión: Tipo Proceso Polvo.

	DIF_HORA	KG	RAP_TON_HORA	TIME_PRE
Media	1.11	23,416.07	22.03	0.23
Mediana	0.77	16,190.04	22.26	0.15
Moda	0.38	7,804.10	22.14	0.10
Varianza	1.08	463,357,045.93	31.02	0.07
Desv Típica	1.04	21,525.73	5.57	0.27

Mínimo	0.03	1,269.04	0.59	0.03
Máximo	8.55	137,707.10	58.63	5.91

Tabla 7. 5: Medidas de Tendencia Central y Medias de Dispersión: Tipo Proceso Pellet.

	DIF_HORA	KG	RAP_TON_HORA	TIME_PRE
Media	0.85	19,757.19	24.76	0.21
Mediana	0.46	10,806.29	25.55	0.14
Moda	0.38	9,831.06	26.57	0.09
Varianza	0.64	303,755,368.07	24.33	0.09
Desv Típica	0.80	17,428.58	4.93	0.30
Mínimo	0.04	1,790.13	2.85	0.04
Máximo	5.23	109,610.10	47.13	6.34

Tabla 7. 6: Medidas de Tendencia Central y Medias de Dispersión: Tipo Proceso Expandido.

Turno Tabla 7. 7 Tabla 7. 8

	DIF_HORA	KG	RAP_TON_HORA	TIME_PRE
Media	0.86	18,667.43	23.00	0.21
Mediana	0.55	12,599.02	23.51	0.14
Moda	0.43	9,495.07	25.24	0.09
Varianza	0.74	310,150,539.50	30.14	0.06
Desv Típica	0.86	17,611.09	5.49	0.25
Mínimo	0.02	1,269.04	0.59	0.04
Máximo	7.78	137,707.10	83.82	7.02

Tabla 7. 7: Medidas de Tendencia Central y Medias de Dispersión: Turno Diurno.

	DIF_HORA	KG	RAP_TON_HORA	TIME_PRE
Media	0.83	18,197.49	23.37	0.24
Mediana	0.54	12,596.83	23.93	0.16
Moda	0.42	9,547.41	25.42	0.10
Varianza	0.66	279,550,528.32	29.36	0.08
Desv Típica	0.81	16,719.76	5.42	0.29

Mínimo	0.03	1,325.04	1.06	0.03
Máximo	8.55	128,549.90	65.29	5.91

Tabla 7. 8: Medidas de Tendencia Central y Medias de Dispersión: Turno Nocturno.

Día de la semana Tabla 7. 9 Tabla 7. 10 Tabla 7. 11 Tabla 7. 12 Tabla 7. 13 Tabla 7. 14 Tabla 7. 15

	DIF_HORA	KG	RAP_TON_HORA	TIME_PRE
Media	0.87	18,855.02	23.05	0.21
Mediana	0.57	12,603.84	23.64	0.14
Moda	0.47	9,607.14	25.34	0.09
Varianza	0.70	294,643,687.85	27.67	0.06
Desv Típica	0.83	17,165.19	5.26	0.25
Mínimo	0.03	1,339.36	2.14	0.04
Máximo	7.13	120,968.80	65.29	4.38

Tabla 7. 9; Medidas de Tendencia Central y Medias de Dispersión: lunes.

	DIF_HORA	KG	RAP_TON_HORA	TIME_PRE
Media	0.95	21,210.91	23.88	0.22
Mediana	0.60	12,625.50	24.34	0.14
Moda	0.46	11,578.34	25.92	0.09
Varianza	0.85	356,412,943.31	26.00	0.07
Desv Típica	0.92	18,878.90	5.10	0.27
Mínimo	0.02	1,346.28	1.76	0.04
Máximo	8.55	137,707.10	83.82	6.34

Tabla 7. 10: Medidas de Tendencia Central y Medias de Dispersión: martes

	DIF_HORA	KG	RAP_TON_HORA	TIME_PRE
Media	0.97	21,610.10	23.73	0.21
Mediana	0.62	12,655.14	24.15	0.14
Moda	0.52	12,493.81	25.54	0.09
Varianza	0.80	345,560,283.97	23.49	0.06
Desv Típica	0.90	18,589.25	4.85	0.24
Mínimo	0.02	1,348.60	2.48	0.04

Máximo	6.77	128,167.20	80.15	5.38
---------------	------	------------	-------	------

Tabla 7. 11: Medidas de Tendencia Central y Medias de Dispersión: miércoles

	DIF_HORA	KG	RAP_TON_HORA	TIME_PRE
Media	0.90	19,838.87	23.46	0.22
Mediana	0.60	12,623.53	23.92	0.15
Moda	0.48	12,863.34	24.50	0.10
Varianza	0.72	296,571,076.97	23.86	0.06
Desv Típica	0.85	17,221.24	4.88	0.24
Mínimo	0.04	1,337.53	1.10	0.04
Máximo	6.80	128,225.90	44.54	3.66

Tabla 7. 12: Medidas de Tendencia Central y Medias de Dispersión: jueves

	DIF_HORA	KG	RAP_TON_HORA	TIME_PRE
Media	0.78	16,835.22	23.00	0.22
Mediana	0.51	12,583.26	23.52	0.15
Moda	0.41	9,287.68	25.16	0.09
Varianza	0.63	254,953,965.24	31.04	0.08
Desv Típica	0.79	15,967.28	5.57	0.27
Mínimo	0.03	1,269.04	1.82	0.03
Máximo	6.88	127,918.20	51.42	7.02

Tabla 7. 13: Medidas de Tendencia Central y Medias de Dispersión: viernes

	DIF_HORA	KG	RAP_TON_HORA	TIME_PRE
Media	0.74	15,295.74	22.25	0.25
Mediana	0.47	10,795.43	22.72	0.15
Moda	0.26	5,616.63	24.53	0.10
Varianza	0.65	268,401,920.23	42.96	0.10
Desv Típica	0.80	16,382.98	6.55	0.31
Mínimo	0.03	1,337.66	0.59	0.04
Máximo	6.12	121,747.70	58.63	5.72

Tabla 7. 14: Medidas de Tendencia Central y Medias de Dispersión: sábado

	DIF_HORA	KG	RAP_TON_HORA	TIME_PRE
Media	0.73	15,498.43	22.71	0.23
Mediana	0.50	12,548.04	23.25	0.15
Moda	0.30	8,860.13	25.06	0.09
Varianza	0.54	221,850,495.04	32.51	0.07
Desv Típica	0.73	14,894.65	5.70	0.26
Mínimo	0.04	1,278.12	1.06	0.04

Máximo	5.87	118,085.70	46.56	5.91
---------------	------	------------	-------	------

Tabla 7. 15: Medidas de Tendencia Central y Medias de Dispersión: domingo

Mes Tabla 7. 16 Tabla 7. 17 Tabla 7. 18 Tabla 7. 19 Tabla 7. 20 Tabla 7. 21 Tabla 7.

22 Tabla 7. 23 Tabla 7. 24 Tabla 7. 25 Tabla 7. 26 Tabla 7. 27

	DIF_HORA	KG	RAP_TON_HORA	TIME_PRE
Media	0.83	18,525.66	23.50	0.23
Mediana	0.52	12,592.70	24.08	0.14
Moda	0.33	9,499.19	25.84	0.09
Varianza	0.75	323,627,375.70	27.82	0.06
Desv Típica	0.87	17,989.65	5.27	0.25
Mínimo	0.03	1,341.16	2.58	0.04
Máximo	7.78	128,030.50	52.77	3.69

Tabla 7. 16: Medidas de Tendencia Central y Medias de Dispersión: Mes Enero.

	DIF_HORA	KG	RAP_TON_HORA	TIME_PRE
Media	0.82	18,081.06	25.56	0.23
Mediana	0.51	12,592.24	24.29	0.14
Moda	0.44	9,750.24	25.51	0.09
Varianza	0.63	274,347,674.04	31.12	0.09
Desv Típica	0.79	16,563.44	5.58	0.30
Mínimo	0.04	1,343.76	2.81	0.04
Máximo	5.26	108,276.40	46.48	4.34

Tabla 7. 17: Medidas de Tendencia Central y Medias de Dispersión: Mes Febrero.

	DIF_HORA	KG	RAP_TON_HORA	TIME_PRE
Media	0.83	18,652.58	23.64	0.21
Mediana	0.54	12,587.52	24.28	0.14
Moda	0.49	9,732.01	25.71	0.09
Varianza	0.64	302,128,009.17	28.06	0.04
Desv Típica	0.80	17,381.83	5.30	0.19
Mínimo	0.04	1,346.28	1.99	0.04
Máximo	5.41	115,153.10	41.59	2.45

Tabla 7. 18: Medidas de Tendencia Central y Medias de Dispersión: Mes Marzo.

	DIF_HORA	KG	RAP_TON_HORA	TIME_PRE
Media	0.85	18,132.54	22,97	0.24
Mediana	0.54	12,579.58	23,17	0.15
Moda	0.46	10,226.03	24,69	0.09
Varianza	0.73	296,430,068.36	30,33	0.09

Desv Típica	0.86	17,217.14	5,51	0.30
Mínimo	0.03	1,346.10	1,10	0.04
Máximo	6.80	112,629.10	53,57	5.91

Tabla 7. 19: Medidas de Tendencia Central y Medias de Dispersión: Mes Abril.

	DIF_HORA	KG	RAP_TON_HORA	TIME_PRE
Media	0.85	18,452.13	23.21	0.23
Mediana	0.54	12,592.76	23.58	0.14
Moda	0.34	9,321.64	24.44	0.10
Varianza	0.72	316,535,609.19	33.97	0.06
Desv Típica	0.85	17,791.45	5.83	0.25
Mínimo	0.02	1,339.16	0.59	0.04
Máximo	5.78	128,225.90	80.15	3.63

Tabla 7. 20: Medidas de Tendencia Central y Medias de Dispersión: Mes Mayo.

	DIF_HORA	KG	RAP_TON_HORA	TIME_PRE
Media	0.84	18,436.41	23.19	0.23
Mediana	0.56	12,610.67	23.60	0.15
Moda	0.46	9,532.68	25.24	0.10
Varianza	0.67	289,851,658.61	28.99	0.09
Desv Típica	0.82	17,025.03	5.38	0.30
Mínimo	0.02	1,337.66	4.45	0.04
Máximo	6.13	126,048.30	83.82	4.47

Tabla 7. 21: Medidas de Tendencia Central y Medias de Dispersión: Mes Junio.

	DIF_HORA	KG	RAP_TON_HORA	TIME_PRE
Media	0.85	18,565.13	23.25	0.24
Mediana	0.54	12,623.24	23.78	0.15
Moda	0.43	9,604.12	25.41	0.10
Varianza	0.67	278,074,094.15	28.50	0.11
Desv Típica	0.82	16,675.55	5.34	0.33
Mínimo	0.03	1,337.53	2.14	0.04
Máximo	6.14	128,549.90	58.63	7.02

Tabla 7. 22: Medidas de Tendencia Central y Medias de Dispersión: Mes Julio.

	DIF_HORA	KG	RAP_TON_HORA	TIME_PRE
Media	0.84	18,042.98	23,06	0.22
Mediana	0.53	12,589.05	23,53	0.15
Moda	0.33	9,387.46	25,19	0.09
Varianza	0.71	289,314,029.77	28,31	0.07

Desv Típica	0.84	17,009.23	5,32	0.27
Mínimo	0.04	1,269.04	1,40	0.06
Máximo	6.29	128,167.20	47,13	6.34

Tabla 7. 23: Medidas de Tendencia Central y Medias de Dispersión: Mes Agosto.

	DIF_HORA	KG	RAP_TON_HORA	TIME_PRE
Media	0.87	18,620.70	22.71	0.23
Mediana	0.57	12,618.46	23.33	0.15
Moda	0.47	9,302.72	25.01	0.10
Varianza	0.74	291,000,657.29	28.89	0.06
Desv Típica	0.86	17,058.74	5.37	0.25
Mínimo	0.03	1,334.96	1.76	0.03
Máximo	8.55	115,597.30	51.42	4.69

Tabla 7. 24: Medidas de Tendencia Central y Medias de Dispersión: Mes Septiembre.

	DIF_HORA	KG	RAP_TON_HORA	TIME_PRE
Media	0.87	18,412.30	22.63	0.22
Mediana	0.55	12,605.63	23.25	0.15
Moda	0.43	9,590.67	25.12	0.09
Varianza	0.71	284,871,981.68	29.99	0.05
Desv Típica	0.84	16,878.15	5.48	0.23
Mínimo	0.04	1,325.04	2.48	0.04
Máximo	6.88	120,968.80	44.54	3.57

Tabla 7. 25: Medidas de Tendencia Central y Medias de Dispersión: Mes Octubre.

	DIF_HORA	KG	RAP_TON_HORA	TIME_PRE
Media	0.87	18,927.04	23.21	0.22
Mediana	0.56	12,590.06	23.83	0.15
Moda	0.42	9,610.82	25.44	0.09
Varianza	0.77	309,175,124.28	29.33	0.05
Desv Típica	0.87	17,583.38	5.42	0.22
Mínimo	0.04	1,336.73	1.82	0.04
Máximo	6.77	127,918.20	42.04	3.09

Tabla 7. 26: Medidas de Tendencia Central y Medias de Dispersión: Mes Noviembre.

	DIF_HORA	KG	RAP_TON_HORA	TIME_PRE
Media	0.85	18,803.82	23.31	0.22
Mediana	0.56	12,597.76	23.92	0.14
Moda	0.43	9,469.98	24.96	0.09
Varianza	0.71	319,391,564.95	32.29	0.05

Desv Típica	0.84	17,871.53	5.68	0.23
Mínimo	0.04	1,341.94	3.16	0.04
Máximo	6.76	137,707.10	46.59	2.35

Tabla 7. 27: Medidas de Tendencia Central y Medias de Dispersión: Mes Diciembre.

Año Tabla 7. 28 Tabla 7. 29 Tabla 7. 30 Tabla 7. 31

	DIF_HORA	KG	RAP_TON_HORA	TIME_PRE
Media	0.89	18,909.90	22.30	0.23
Mediana	0.63	12,621.61	22.83	0.14
Moda	0.54	12,332.78	23.85	0.09
Varianza	0.70	289,815,354.79	26.01	0.07
Desv Típica	0.84	17,023.96	5.10	0.26
Mínimo	0.03	1,269.04	1.83	0.04
Máximo	6.88	127,918.20	46.04	4.10

Tabla 7. 28: Medidas de Tendencia Central y Medias de Dispersión: Año 2015.

	DIF_HORA	KG	RAP_TON_HORA	TIME_PRE
Media	0.81	17,785.43	23.32	0.23
Mediana	0.53	12,584.24	23.90	0.14
Moda	0.46	9,508.97	25.43	0.09
Varianza	0.61	268,891,658.77	31.44	0.08
Desv Típica	0.78	16,397.92	5.61	0.28
Mínimo	0.02	1,336.46	1.10	0.04
Máximo	6.80	137,707.10	83.82	6.34

Tabla 7. 29: Medidas de Tendencia Central y Medias de Dispersión: Año 2016.

	DIF_HORA	KG	RAP_TON_HORA	TIME_PRE
Media	0.84	18,363.71	23.68	0.21
Mediana	0.51	12,572.66	24.27	0.14
Moda	0.43	9,516.19	25.45	0.09
Varianza	0.72	295,728,959.38	32.77	0.07
Desv Típica	0.85	17,196.77	5.72	0.26
Mínimo	0.03	1,278.12	0.59	0.03
Máximo	7.78	128,030.50	51.42	7.02

Tabla 7. 30: Medidas de Tendencia Central y Medias de Dispersión: Año 2017.

	DIF_HORA	KG	RAP_TON_HORA	TIME_PRE
Media	0.88	19,076.76	23.02	0.23
Mediana	0.55	12,627.88	23.58	0.16
Moda	0.34	10,838.13	24.50	0.10
Varianza	0.79	338,566,564.50	26.24	0.06

Desv Típica	0.89	18,400.18	5.12	0.25
Mínimo	0.04	1,346.79	1.06	0.04
Máximo	8.55	128,549.90	50.13	5.72

Tabla 7. 31: Medidas de Tendencia Central y Medias de Dispersión: Año 2018.

7.2. Apéndice B: Gráfico de barras de Análisis Univariante Categórico

Presentación Ilustración 7: 1;Error! No se encuentra el origen de la referencia.

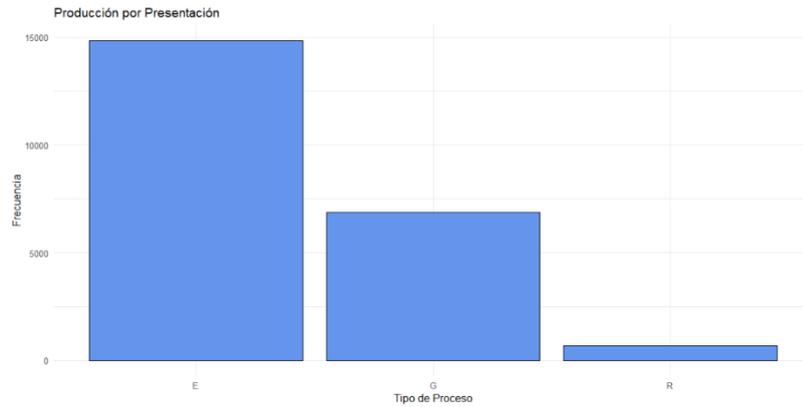


Ilustración 7: 1: Gráfico de barras Presentación.

Tipo de Proceso Ilustración 7: 2 Ilustración 7: 2

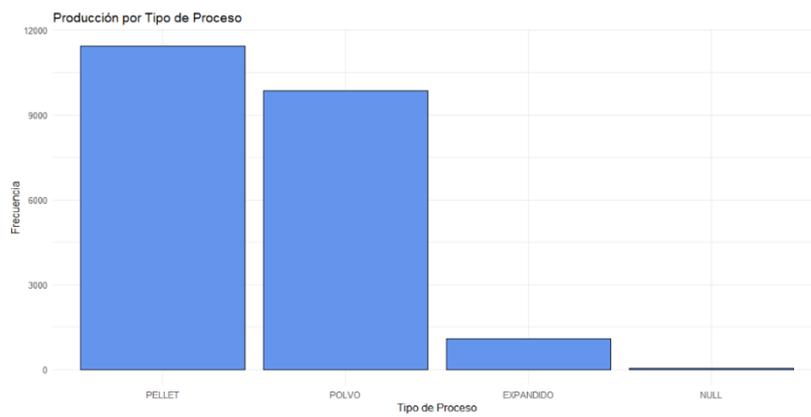


Ilustración 7: 2: Gráfico de barras Tipo de Proceso.

Turno Ilustración 7: 3

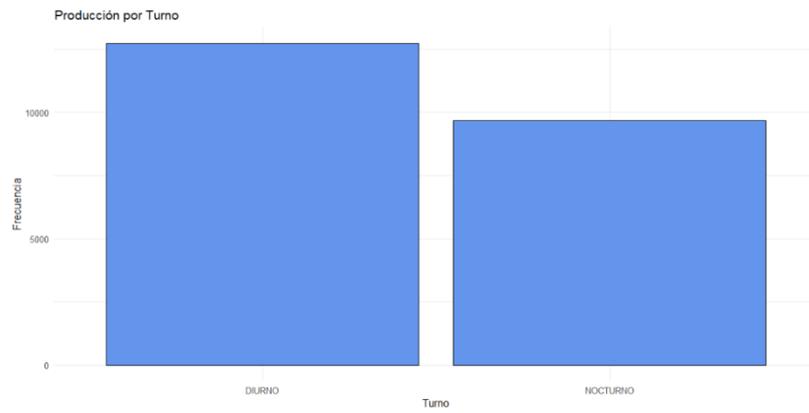


Ilustración 7: 3: Gráfico de barras Turno.

Día Ilustración 7: 4

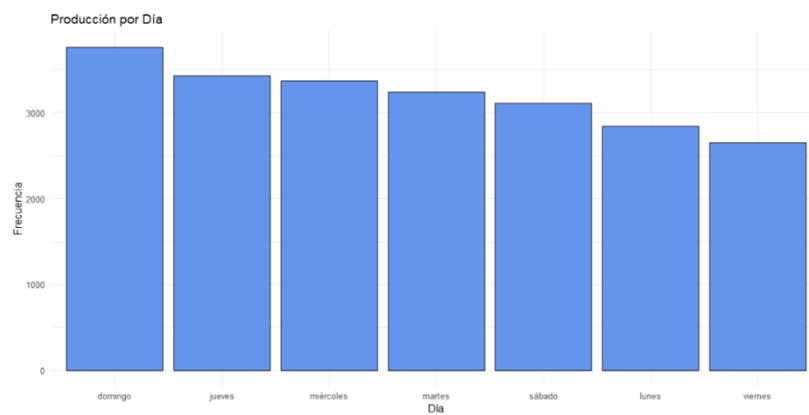


Ilustración 7: 4: Gráfico de barras Día

Mes Ilustración 7: 5

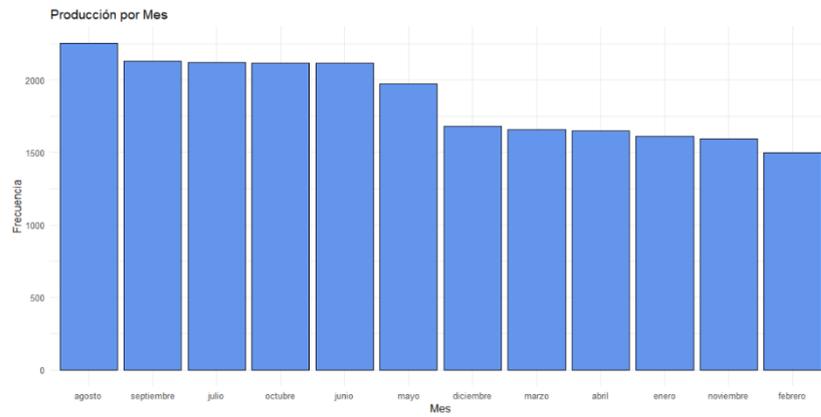


Ilustración 7: 5: Gráfico de barras Mes

Año Ilustración 7: 6

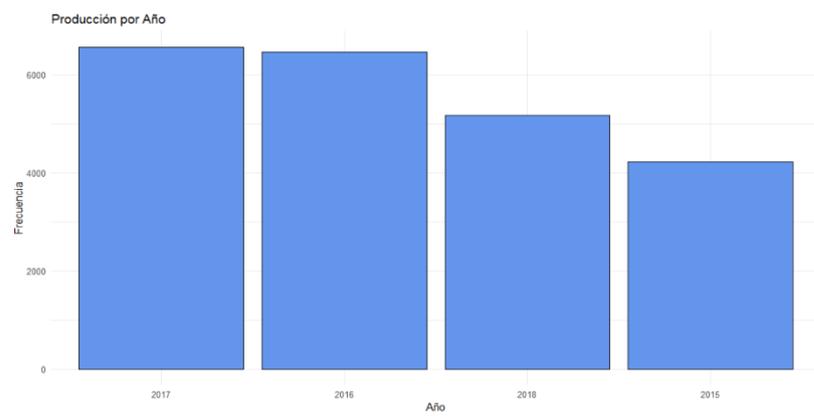


Ilustración 7: 6: Gráfico de barras Año

7.3. Apéndice B: Diagrama de Cajas de Análisis Bivariante: Univariante Cuantitativo Univariante vs Multivariante Categórico.

Variables cuantitativas vs Presentación Ilustración 7: 7 Ilustración 7: 8 Ilustración 7: 9 Ilustración 7: 10

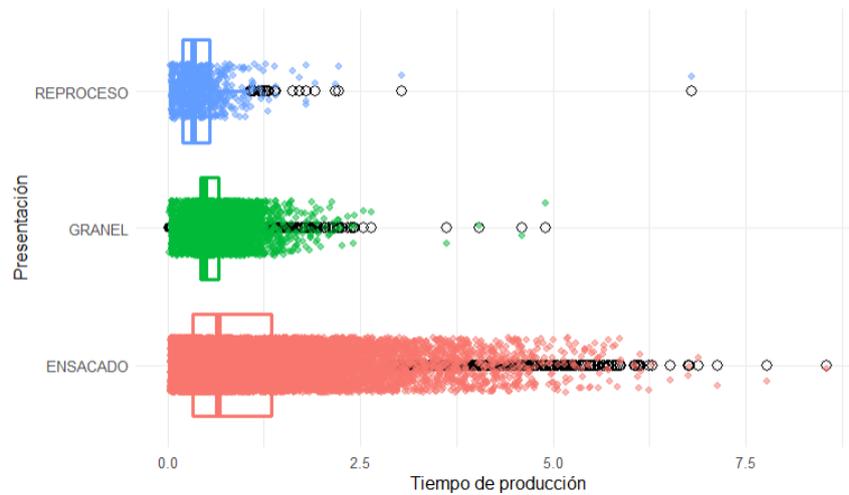


Ilustración 7: 7: Diagrama de Cajas Tiempo de Producción vs Presentación.

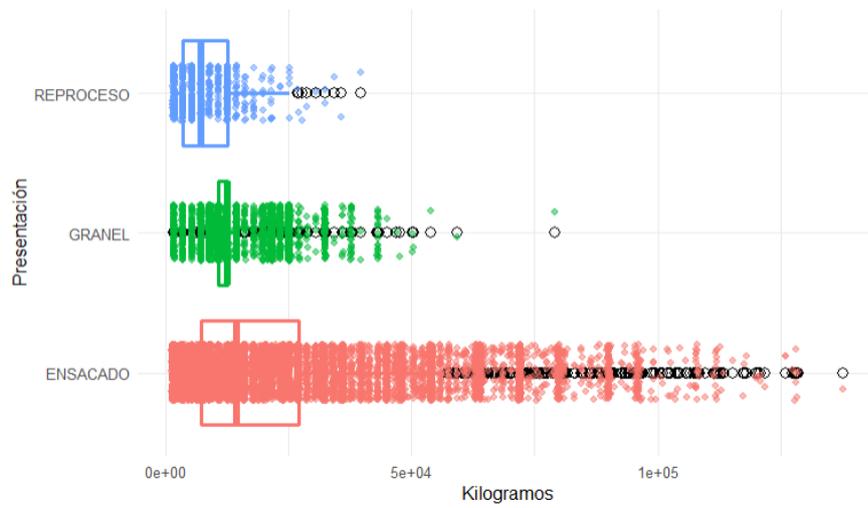


Ilustración 7: 8: Diagrama de Cajas Kilogramos vs Presentación.

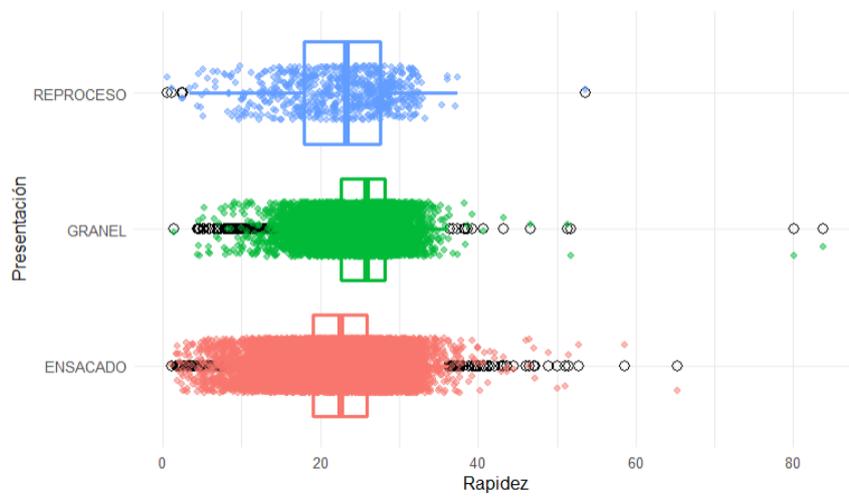


Ilustración 7: 9: Diagrama de Cajas Rapidez vs Presentación.

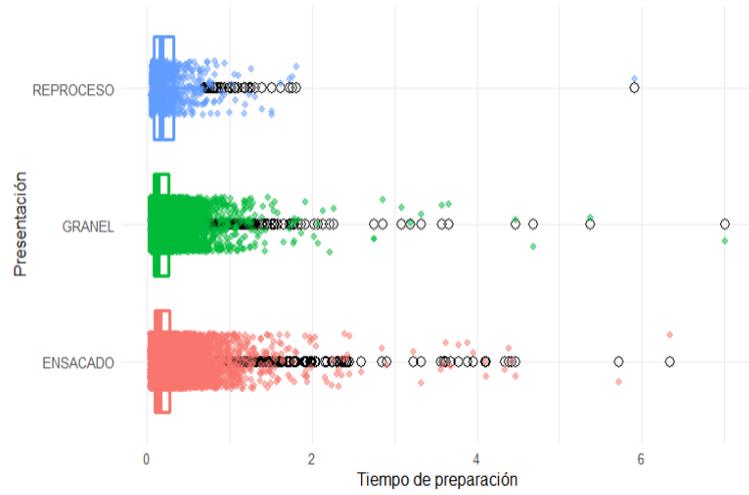


Ilustración 7: 10: Diagrama de Cajas Tiempo Preparación vs Presentación.

Variables Cuantitativas vs Tipo de Proceso Ilustración 7: 11 Ilustración 7: 12

Ilustración 7: 13 Ilustración 7: 14

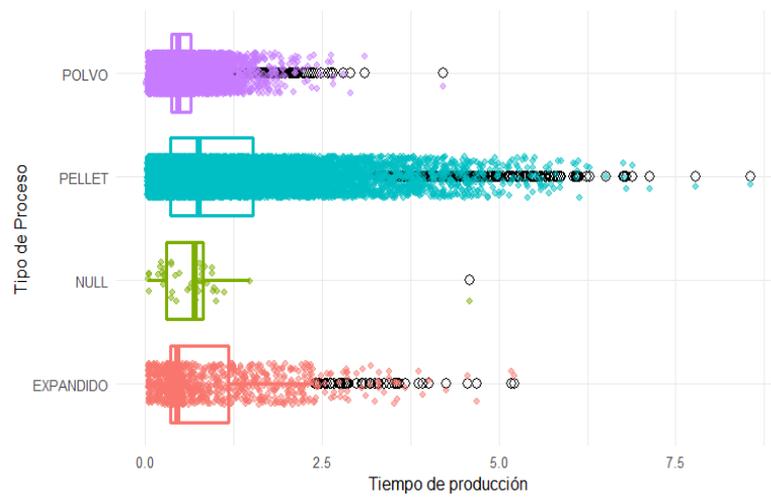


Ilustración 7: 11: Diagrama de Cajas Tiempo de Producción vs Tipo de Proceso.

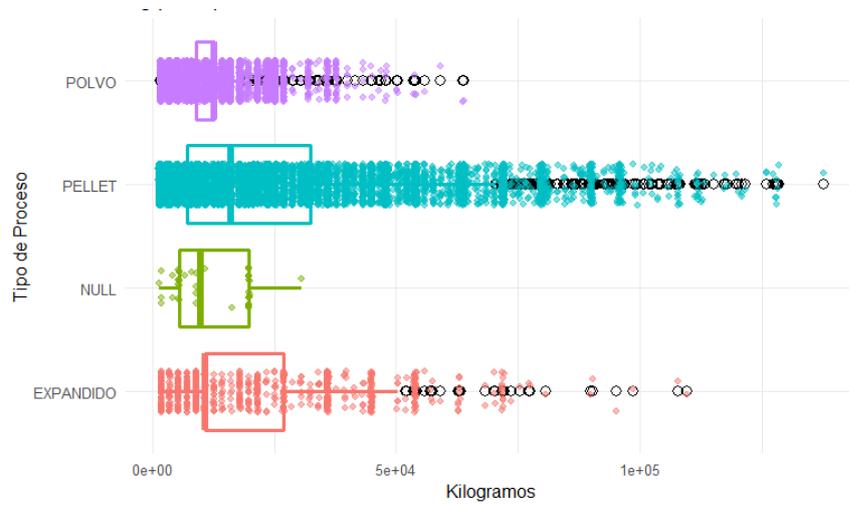


Ilustración 7: 12: Diagrama de cajas Kilogramos vs Tipo de Proceso

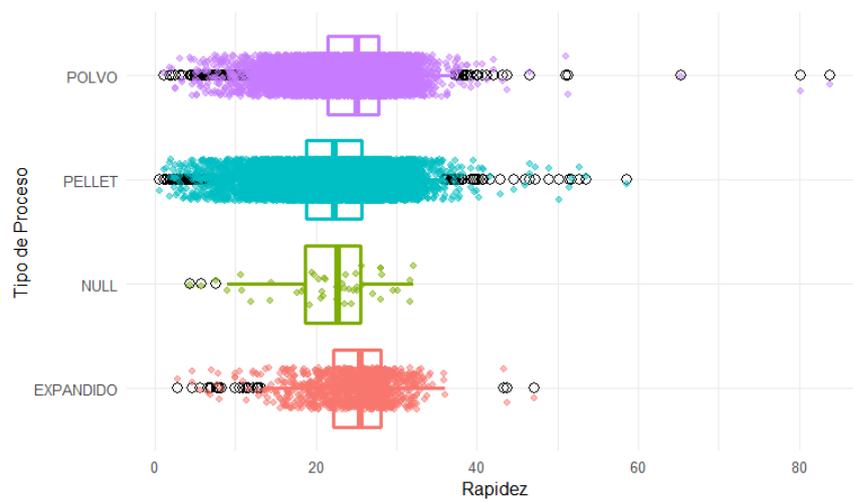


Ilustración 7: 13: Diagrama de cajas Rapidez vs Tipo de Proceso.

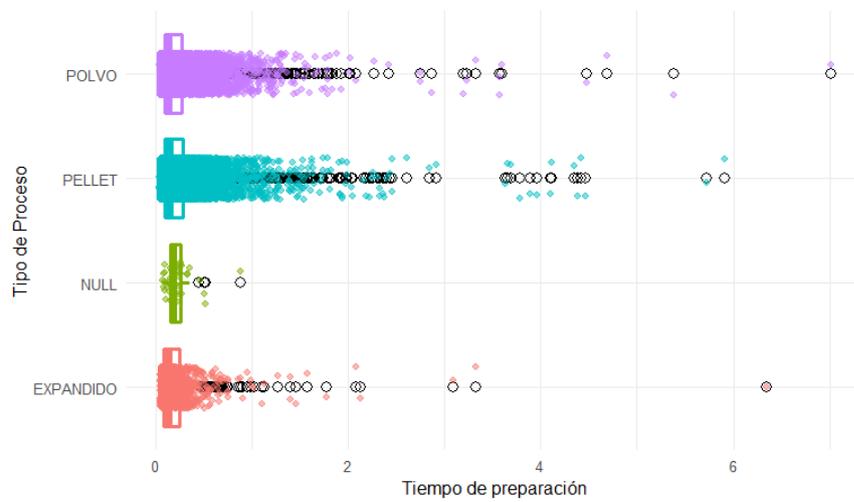


Ilustración 7: 14: Diagrama de cajas Tiempo de Preparación vs Tipo de Proceso

Variables Cuantitativas vs Turno Ilustración 7: 15 Ilustración 7: 16 Ilustración 7: 17

Ilustración 7: 18

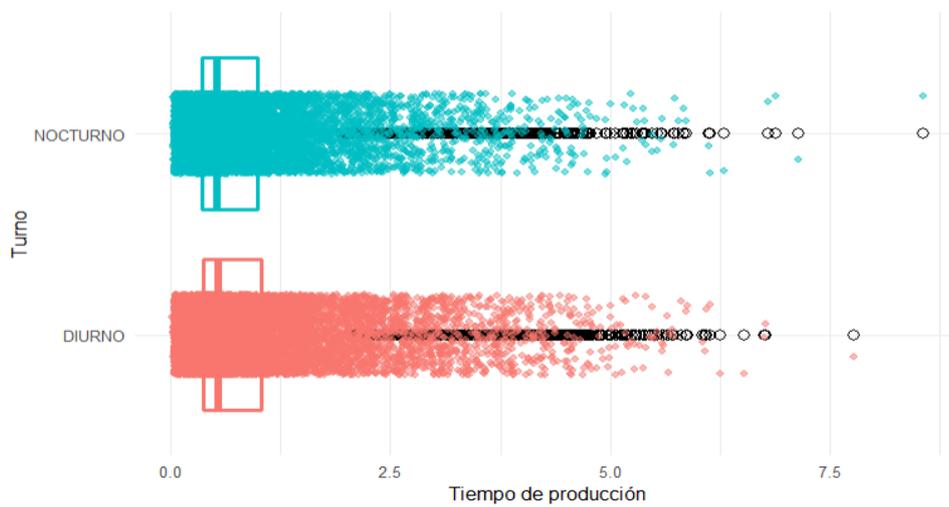


Ilustración 7: 15: Diagrama de cajas Tiempo Producción vs Turno.

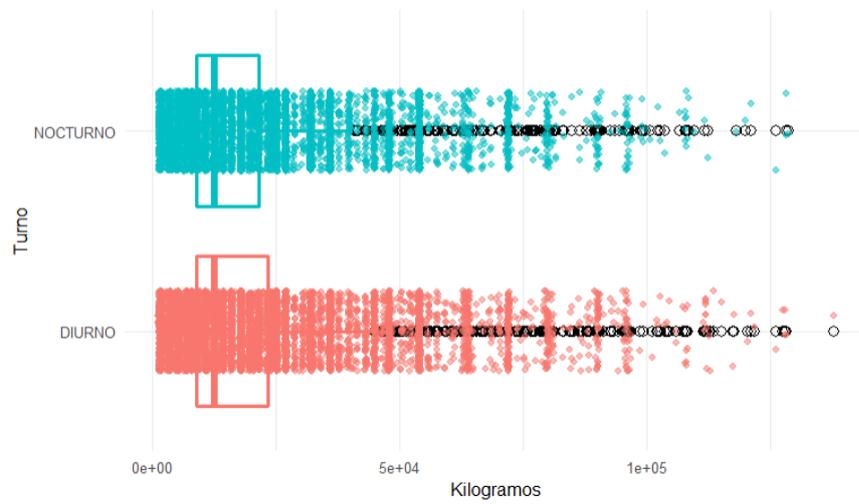


Ilustración 7: 16: Diagrama de cajas Kilogramos vs Turno.

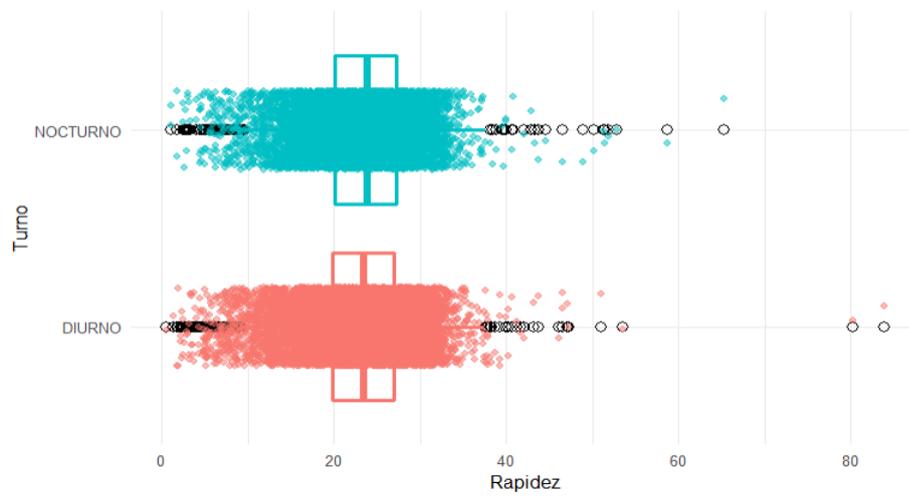


Ilustración 7: 17: Diagrama de cajas Rapidez vs Turno.

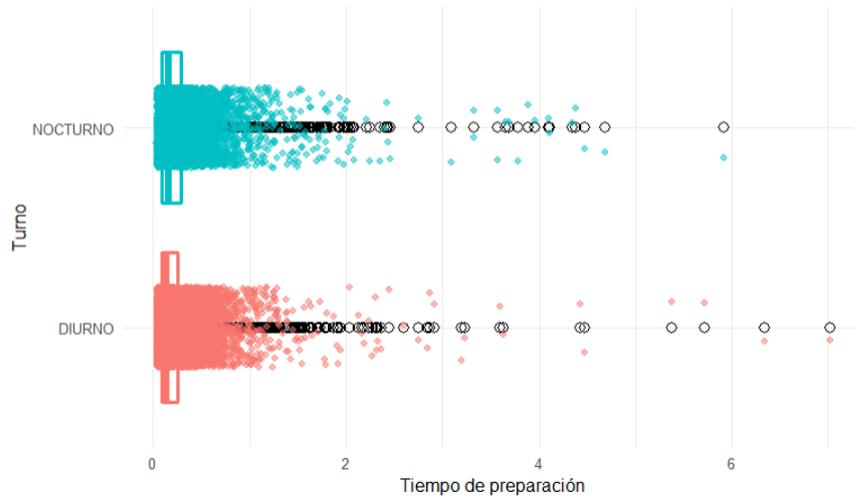


Ilustración 7: 18: Diagrama de cajas Tiempo de Preparación vs Turno.

Variables Cuantitativas vs Día Ilustración 7: 19 Ilustración 7: 20 Ilustración 7: 21

Ilustración 7: 22

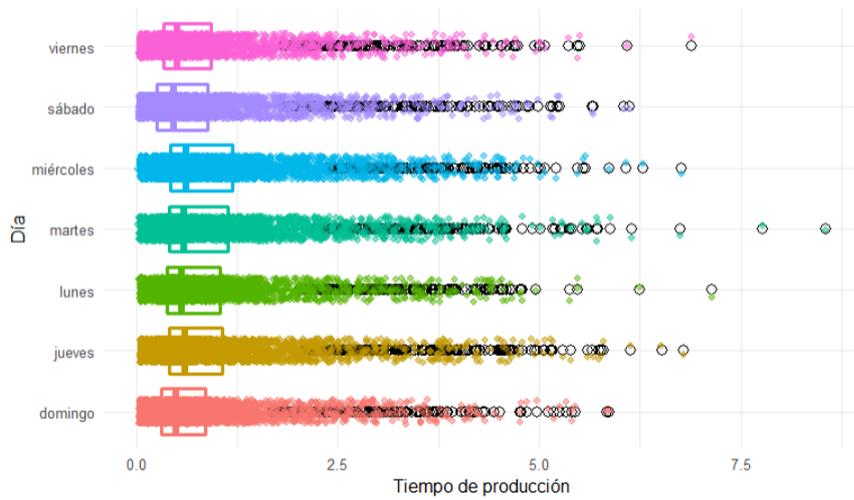


Ilustración 7: 19: Diagrama de cajas Tiempo de Producción vs Día.

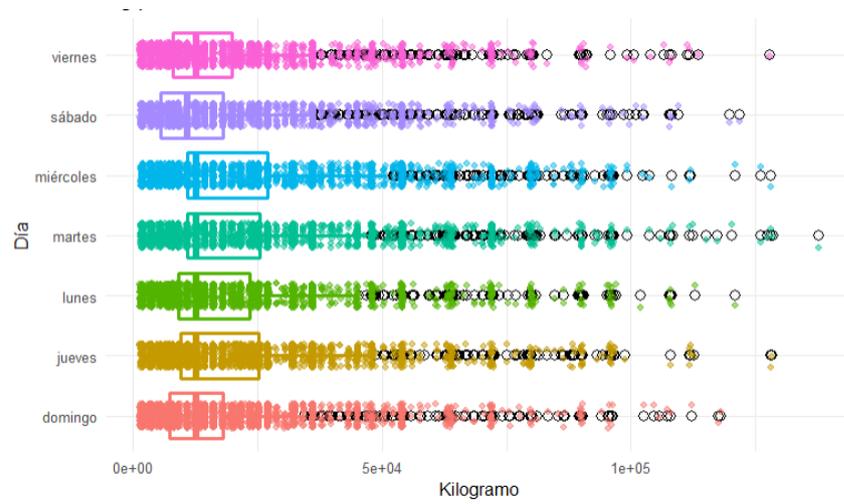


Ilustración 7: 20: Diagrama de cajas Kilogramos vs Día.

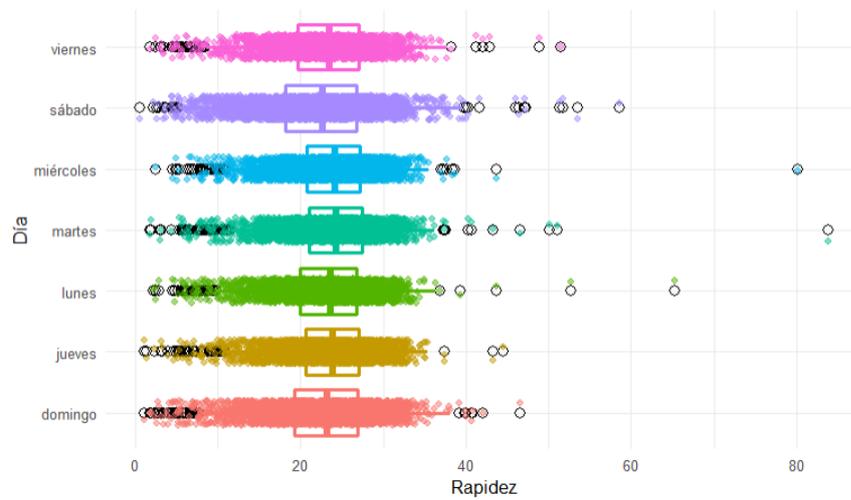


Ilustración 7: 21: Diagrama de cajas Rapidez vs Día.

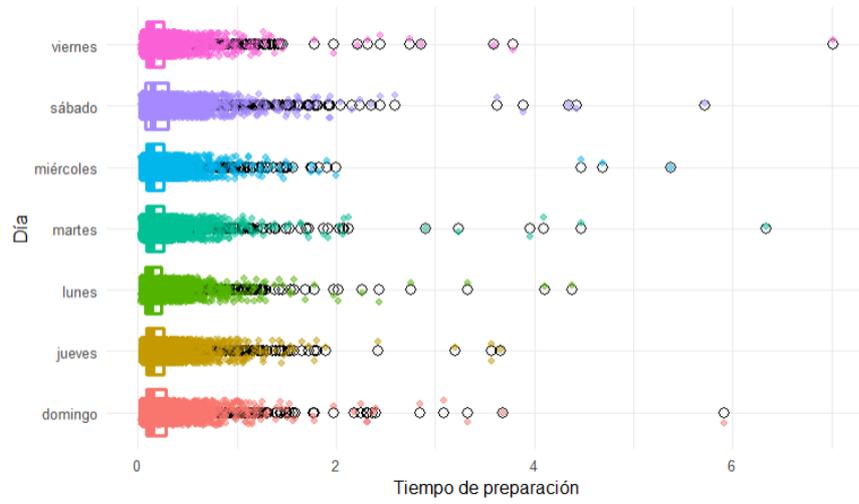


Ilustración 7: 22: Diagrama de cajas Tiempo Preparación vs Día

Variables Cuantitativas vs Mes Ilustración 7: 23 Ilustración 7: 24 Ilustración 7: 25 Ilustración 7: 26

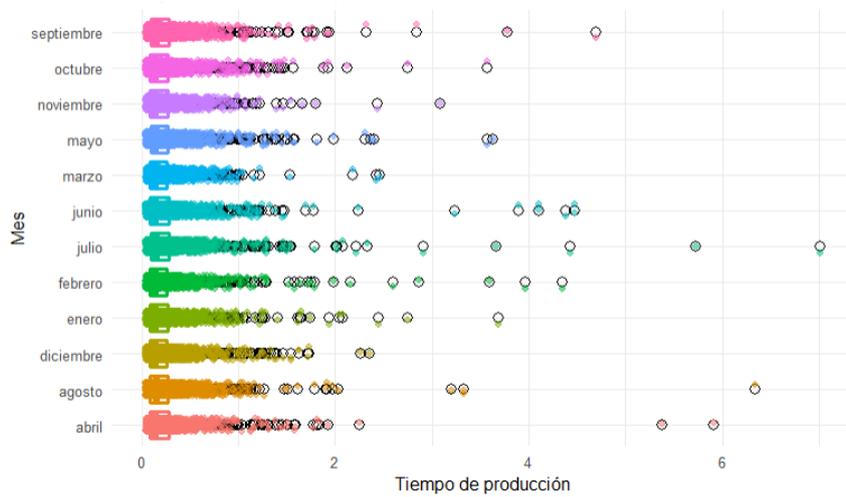


Ilustración 7: 23: Diagrama de cajas Tiempo Producción vs Mes

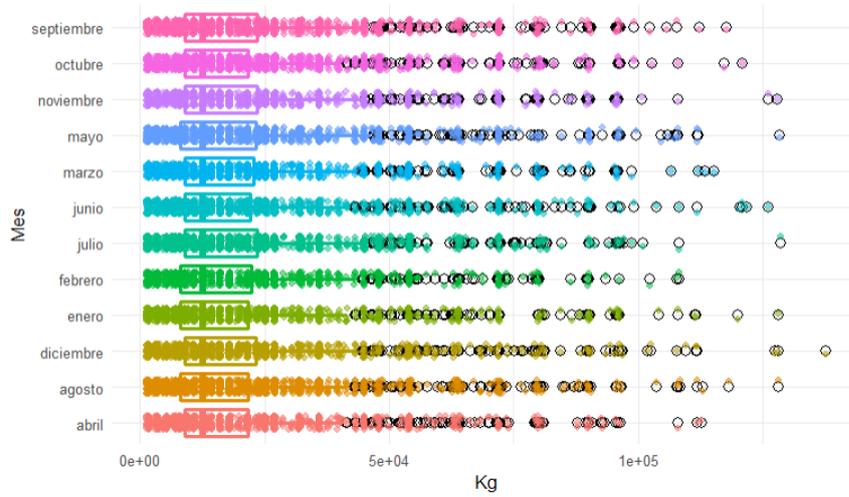


Ilustración 7: 24: Diagrama de cajas Kilogramos vs Mes.

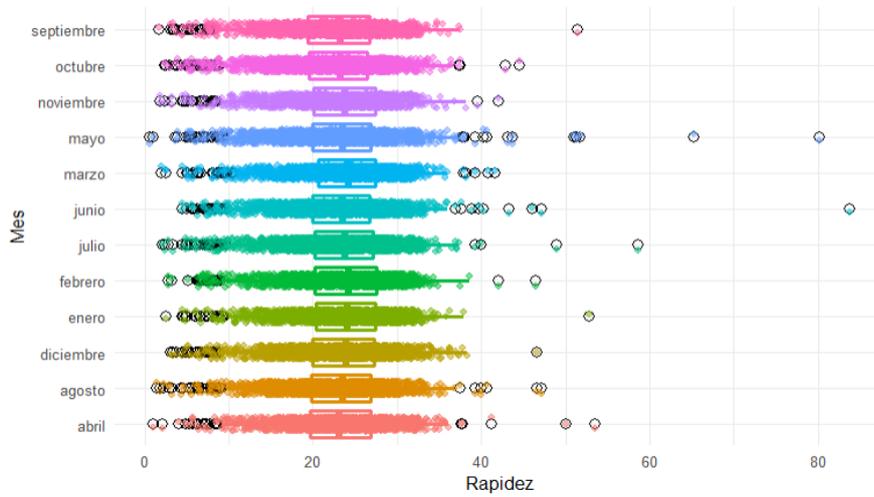


Ilustración 7: 25: Diagrama de cajas Rapidez vs Mes.

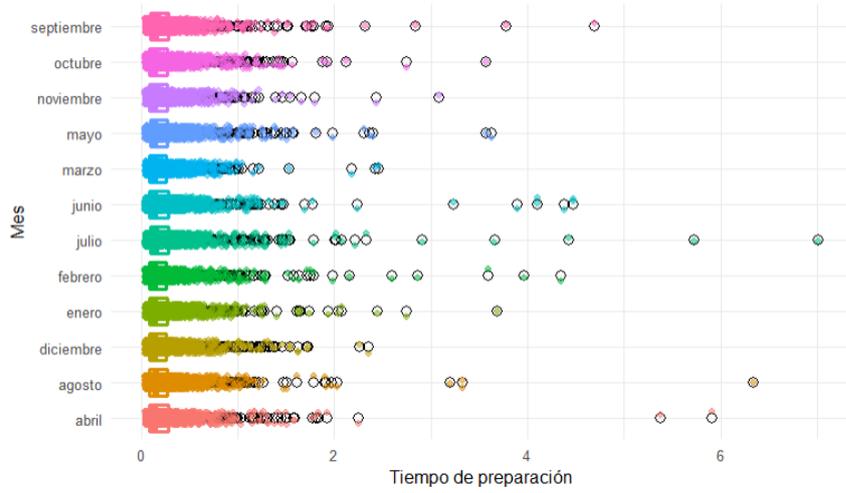


Ilustración 7: 26: Diagrama de cajas Tiempo Preparación vs Mes.

Variables Cuantitativas vs Años Ilustración 7: 27 Ilustración 7: 28 Ilustración 7: 29
 Ilustración 7: 30

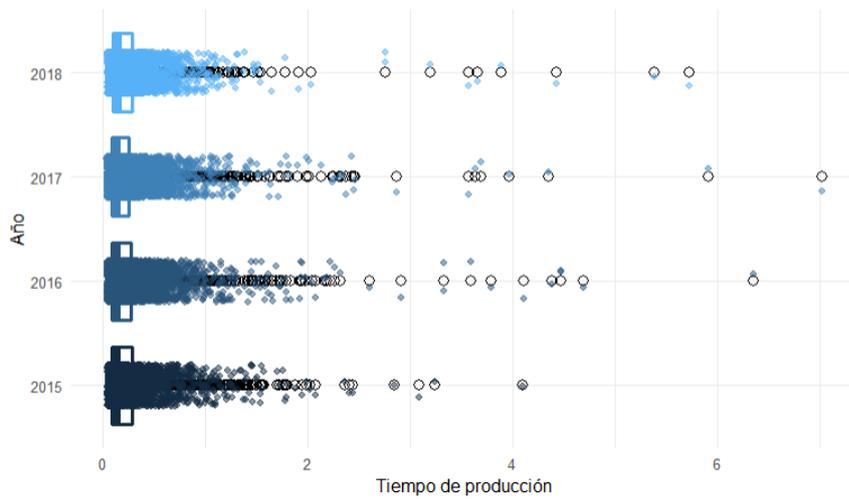


Ilustración 7: 27: Diagrama de cajas Tiempo Producción vs Año.

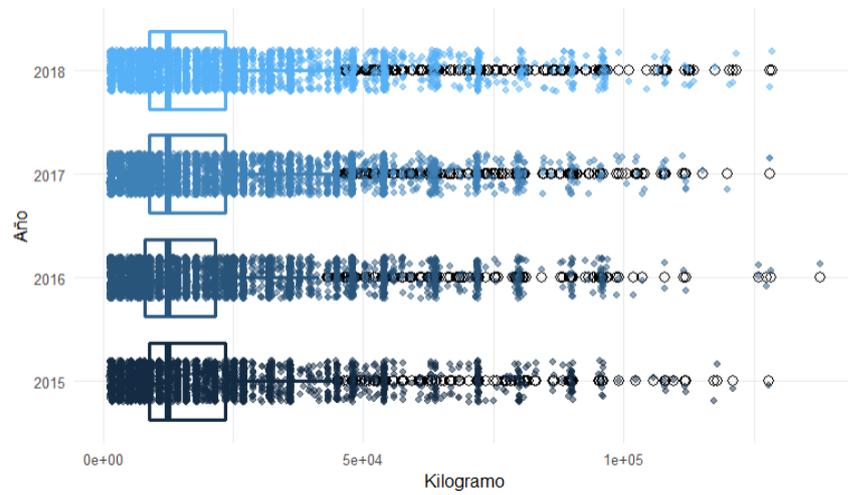


Ilustración 7: 28: Diagrama de cajas Kilogramos vs Año.

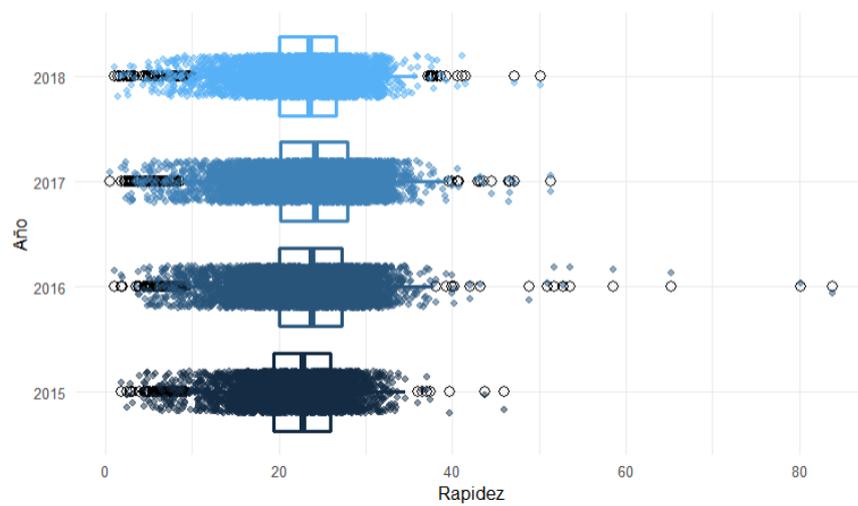


Ilustración 7: 29: Diagrama de cajas Rapidez vs Año.

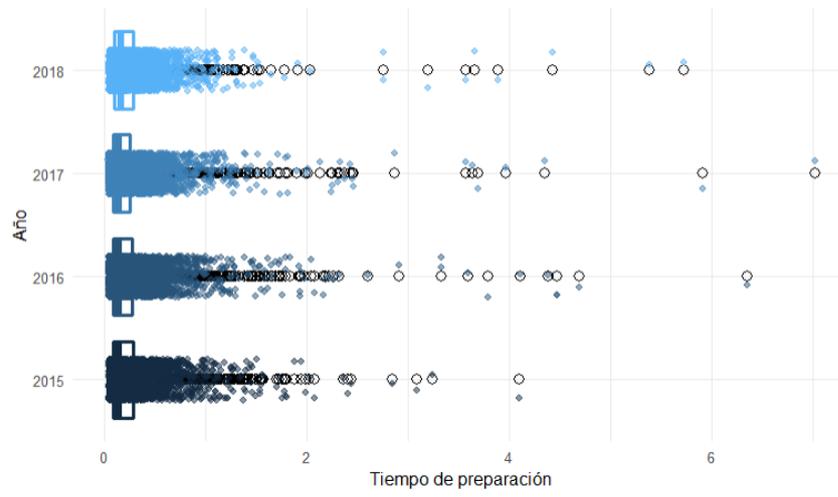


Ilustración 7: 30: Diagrama de cajas Tiempo Preparación vs Año.

7.4. Apéndice D: Gráfico de Dispersión de Análisis Bivariante: Bivariante Cuantitativo Univariante vs Multivariante Categórico.

Tiempo de Producción vs Kilogramos vs Presentación Ilustración 7: 31

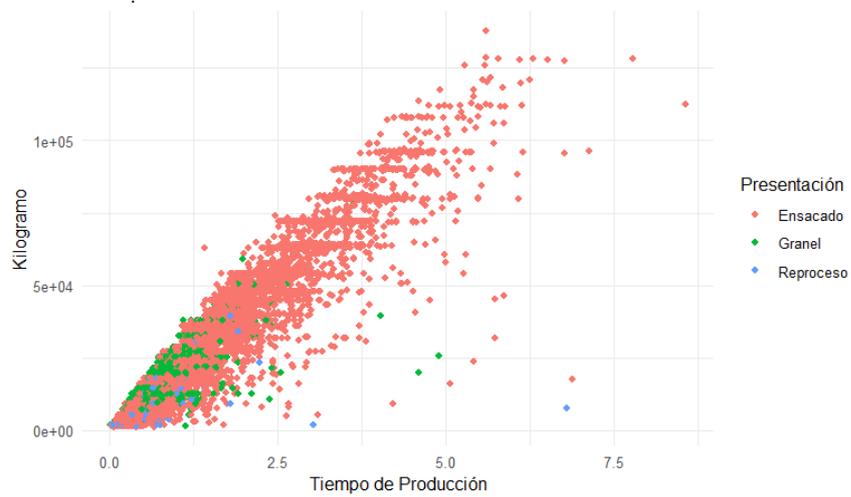


Ilustración 7: 31: Gráfico de Dispersión Tiempo de Producción vs Kilogramo vs Presentación.

Tiempo de Producción vs Rapidez vs Presentación Ilustración 7: 32

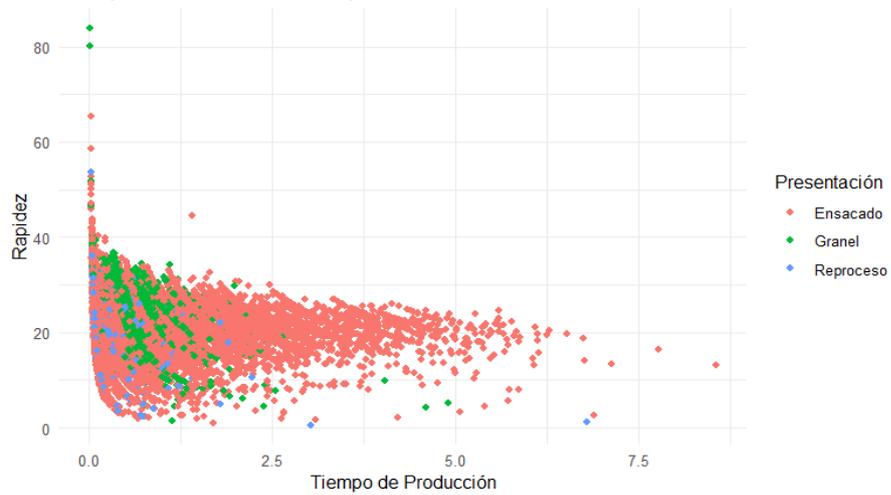


Ilustración 7: 32: Gráfico de Dispersión Tiempo de Producción vs Rapidez vs Presentación.

Tiempo de Producción vs Tiempo de Preparación vs Presentación Ilustración 7:

33

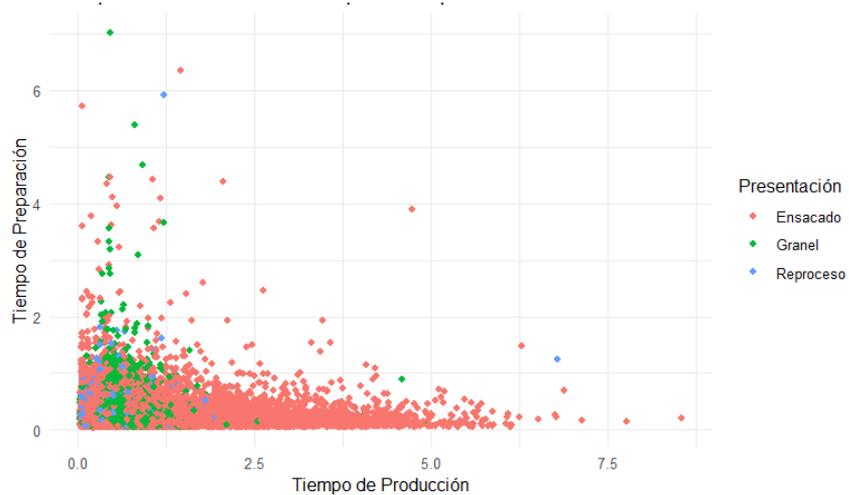


Ilustración 7: 33: Gráfico de Dispersión Tiempo de Producción vs Tiempo de Preparación vs Presentación.

Kilogramos vs Rapidez vs Presentación Ilustración 7: 34

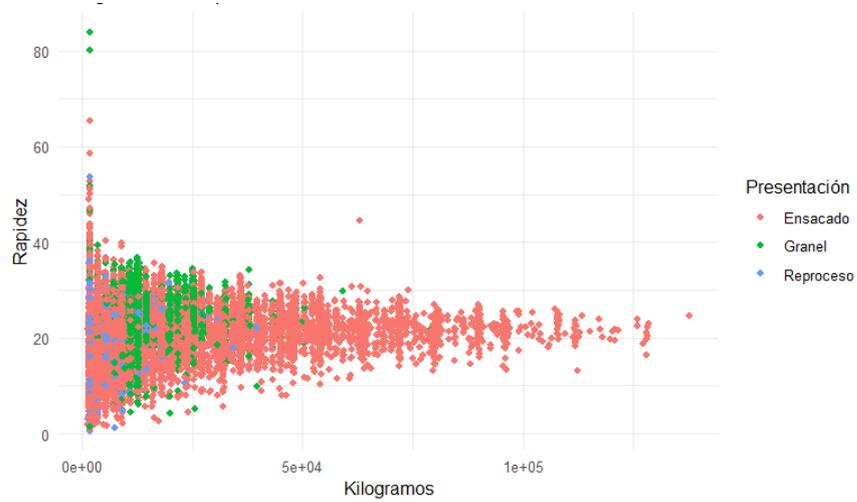


Ilustración 7: 34: Gráfico de Dispersión Kilogramos vs Rapidez vs Presentación.

Kilogramos vs Tiempo de Preparación vs Presentación Ilustración 7: 35

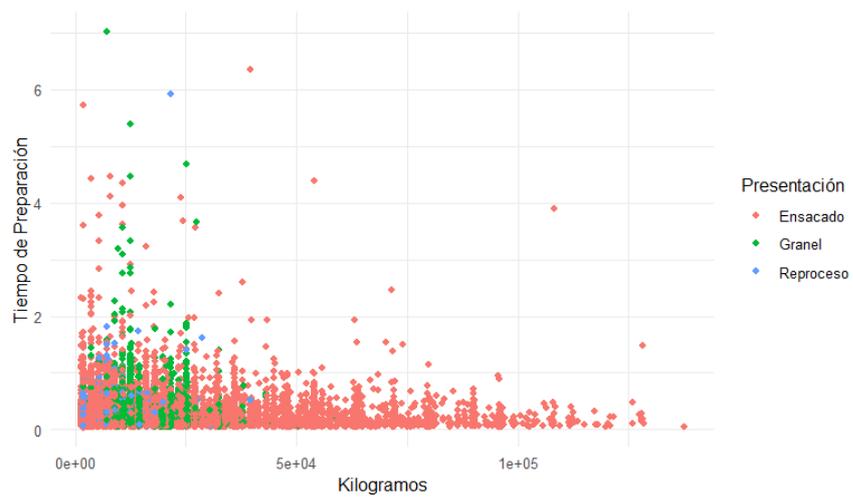


Ilustración 7: 35: Gráfico de Dispersión Kilogramos vs Tiempo de Preparación vs Presentación.

Rapidez vs Tiempo de preparación vs Presentación Ilustración 7: 36

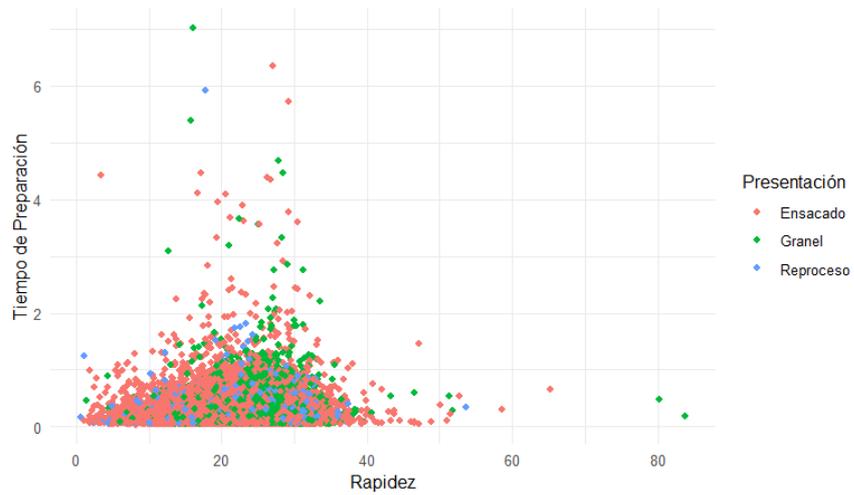


Ilustración 7: 36: Gráfico de Dispersión Rapidez vs Tiempo de Preparación vs Presentación.

Tiempo de Producción vs Kilogramos vs Tipo de Proceso Ilustración 7: 37

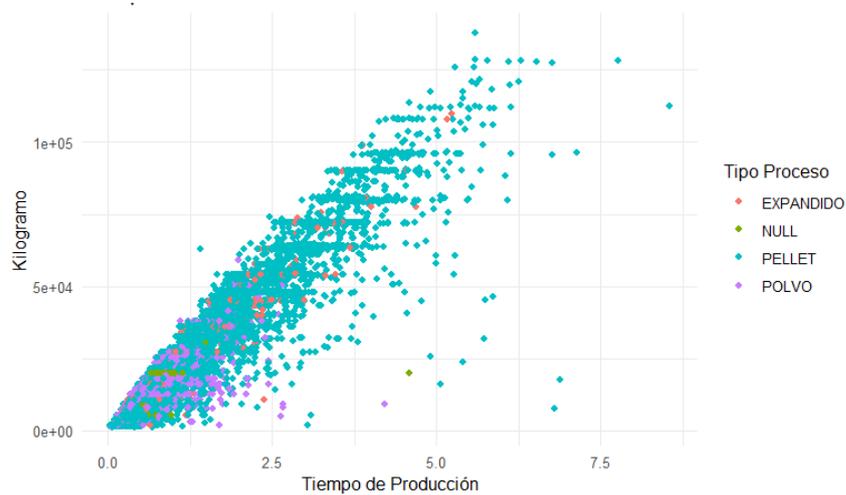


Ilustración 7: 37: Gráfico de Dispersión Tiempo de Producción vs Kilogramo vs Tipo de Proceso.

Tiempo de Producción vs Rapidez vs Tipo de Proceso Ilustración 7: 38

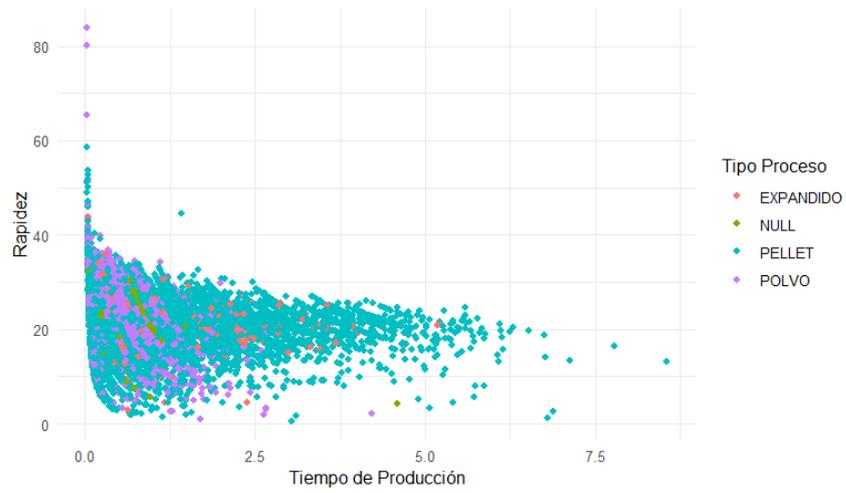


Ilustración 7: 38: Gráfico de Dispersión Tiempo de Producción vs Rapidez vs Tipo de Proceso.

Tiempo de Producción vs Tiempo de Preparación vs Tipo de Proceso Ilustración

7: 39

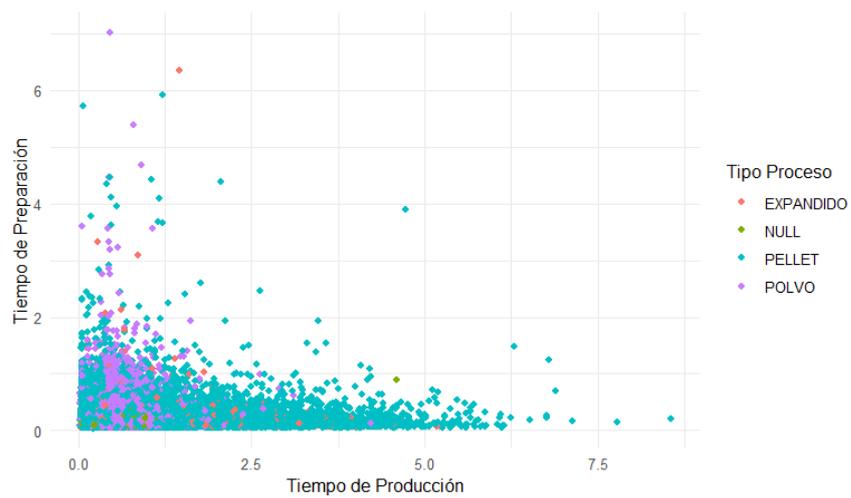


Ilustración 7: 39: Gráfico de Dispersión Tiempo de Producción vs Tiempo de Preparación vs Tipo de Proceso.

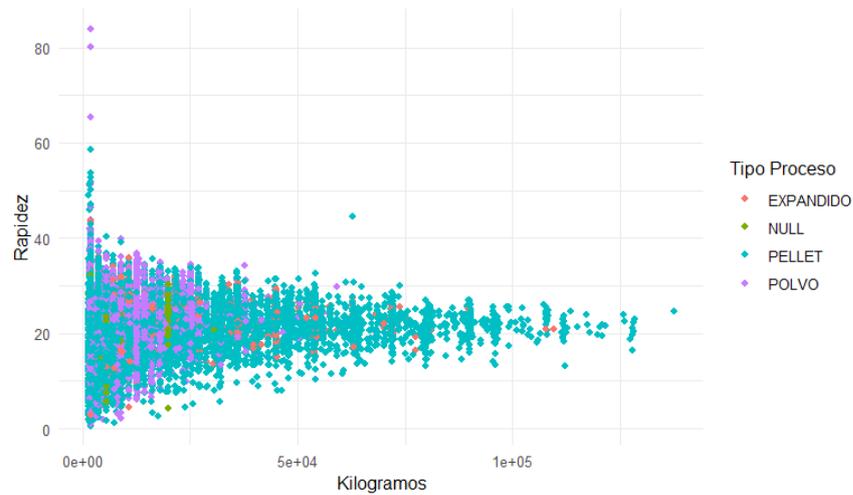
Kilogramos vs Rapidez vs Tipo de Proceso Ilustración 7: 40

Ilustración 7: 40: Gráfico de Dispersión Kilogramos vs Rapidez vs Tipo de Proceso.

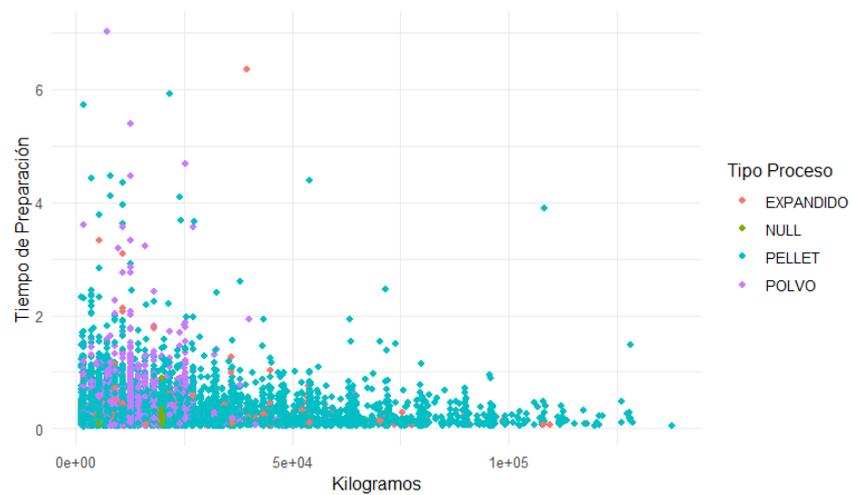
Kilogramos vs Tiempo de Preparación vs Tipo de Proceso Ilustración 7: 41

Ilustración 7: 41: Gráfico de Dispersión Kilogramos vs Tiempo de Preparación vs Tipo de Proceso.

Rapidez vs Tiempo de preparación vs Tipo de Proceso Ilustración 7: 42

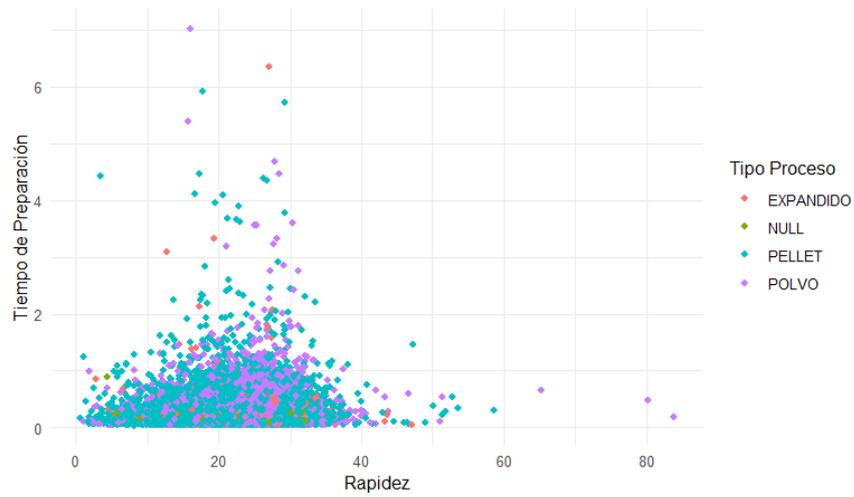


Ilustración 7: 42: Gráfico de Dispersión Rapidez vs Tiempo de Preparación vs Tipo de Proceso.

Tiempo de Producción vs Kilogramos vs Turno Ilustración 7: 43

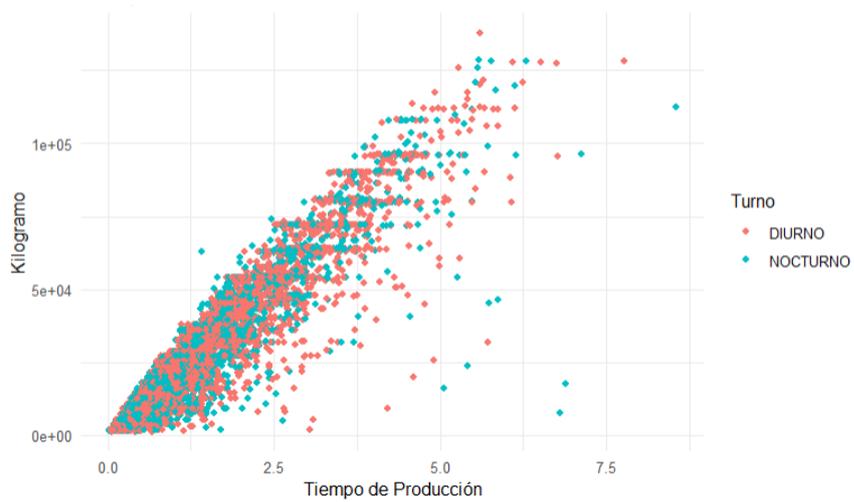


Ilustración 7: 43: Gráfico de Dispersión Tiempo de Producción vs Kilogramos vs Turno.

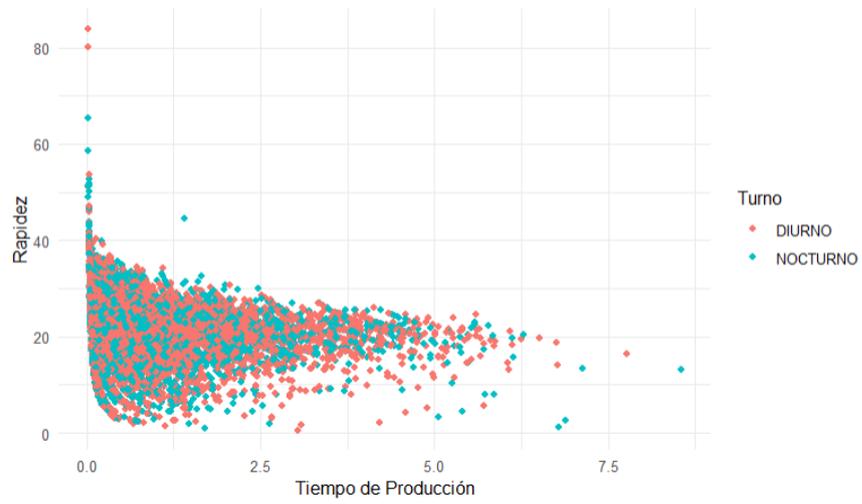
Tiempo de Producción vs Rapidez vs Turno Ilustración 7: 44

Ilustración 7: 44: Gráfico de Dispersión Tiempo de Producción vs Rapidez vs Turno.

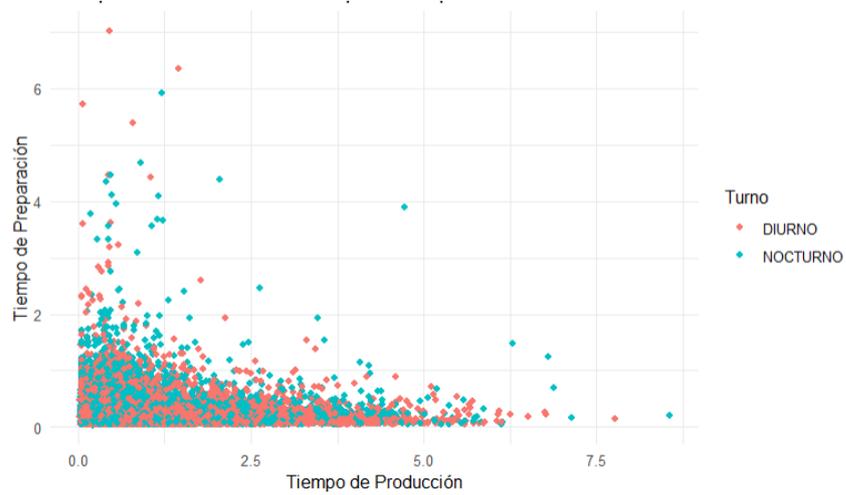
Tiempo de Producción vs Tiempo de Preparación vs Turno Ilustración 7: 45

Ilustración 7: 45: Gráfico de Dispersión Tiempo de Producción vs Tiempo de Preparación vs Turno.

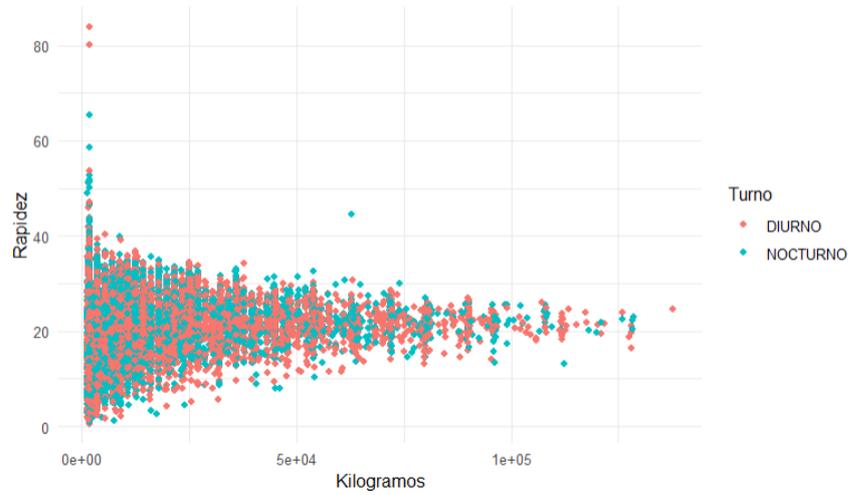
Kilogramos vs Rapidez vs Turno Ilustración 7: 46

Ilustración 7: 46: Gráfico de Dispersión Kilogramos vs Rapidez vs Turno.

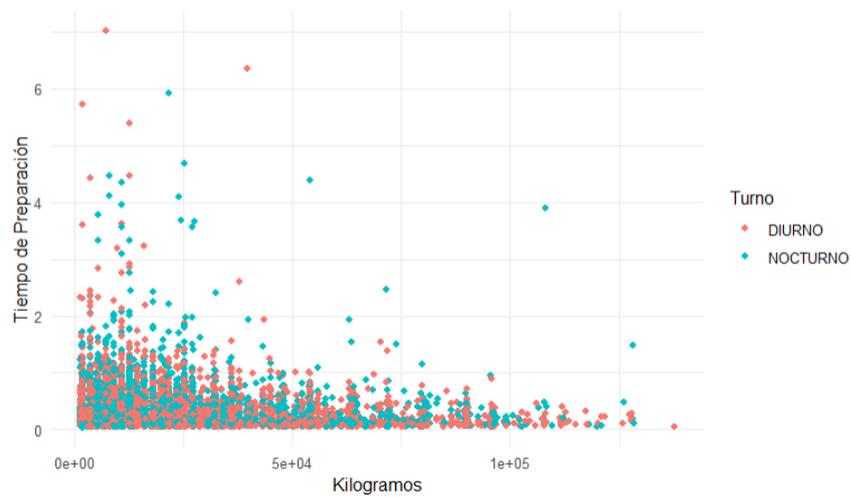
Kilogramos vs Tiempo de Preparación vs Turno Ilustración 7: 47

Ilustración 7: 47: Gráfico de Dispersión Kilogramos vs Tiempo de Preparación vs Turno.

Rapidez vs Tiempo de preparación vs Turno Ilustración 7: 48

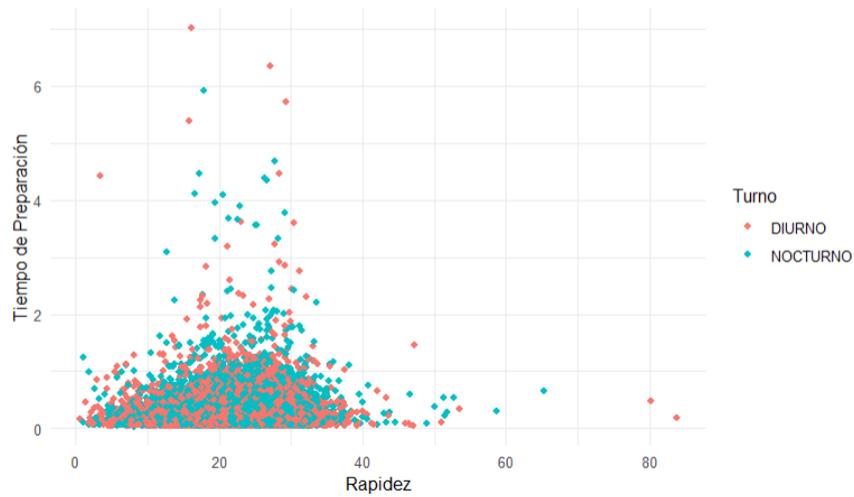


Ilustración 7: 48: Gráfico de Dispersión Rapidez vs Tiempo de Preparación vs Turno.

Tiempo de Producción vs Kilogramos vs Día Ilustración 7: 49

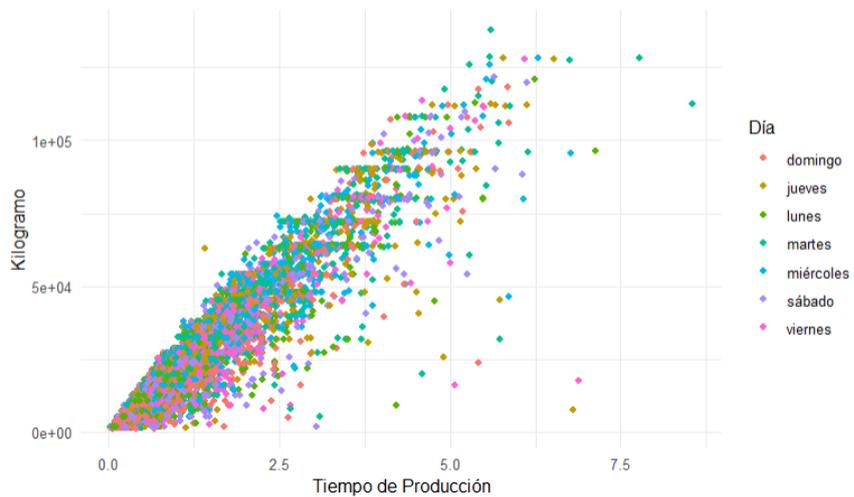


Ilustración 7: 49: Gráfico de Dispersión Tiempo de Producción vs Kilogramos vs Día.

Tiempo de Producción vs Rapidez vs Día Ilustración 7: 50

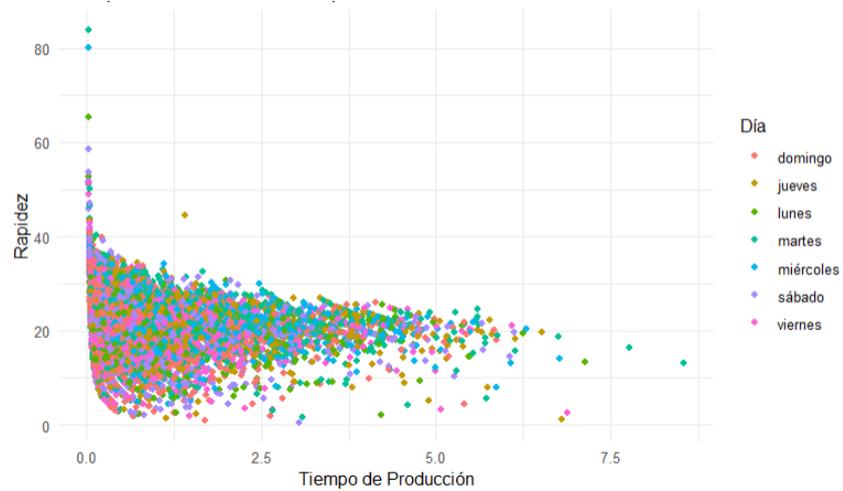


Ilustración 7: 50: Gráfico de Dispersión Tiempo de Producción vs Rapidez vs Día.

Tiempo de Producción vs Tiempo de Preparación vs Día Ilustración 7: 51

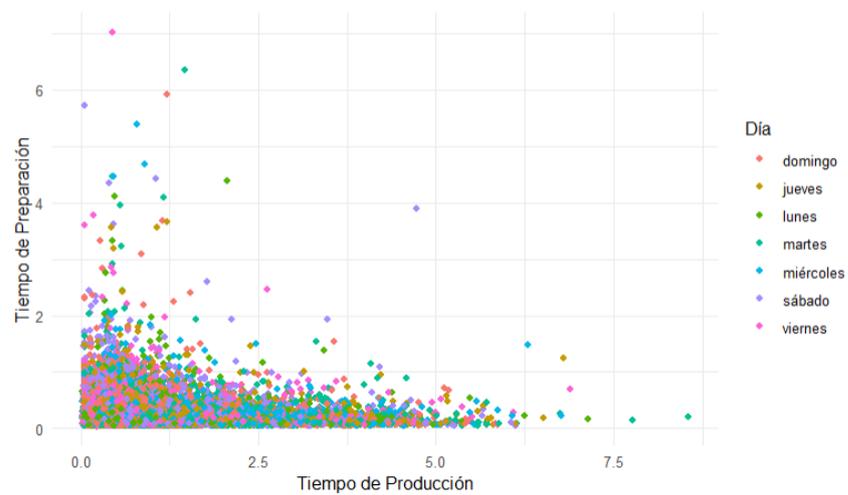


Ilustración 7: 51: Gráfico de Dispersión Tiempo de Producción vs Tiempo de Preparación vs Día.

Kilogramos vs Rapidez vs Día Ilustración 7: 52

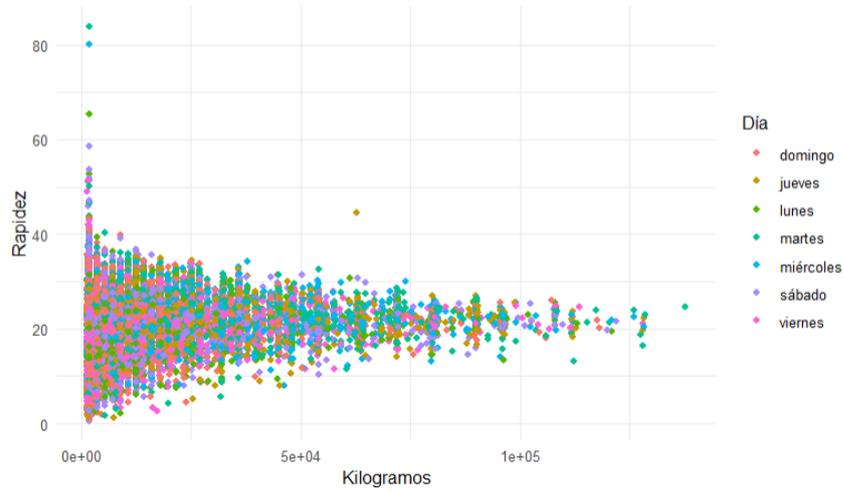


Ilustración 7: 52: Gráfico de Dispersión Kilogramos vs Rapidez vs Día.

Kilogramos vs Tiempo de Preparación vs Día Ilustración 7: 53

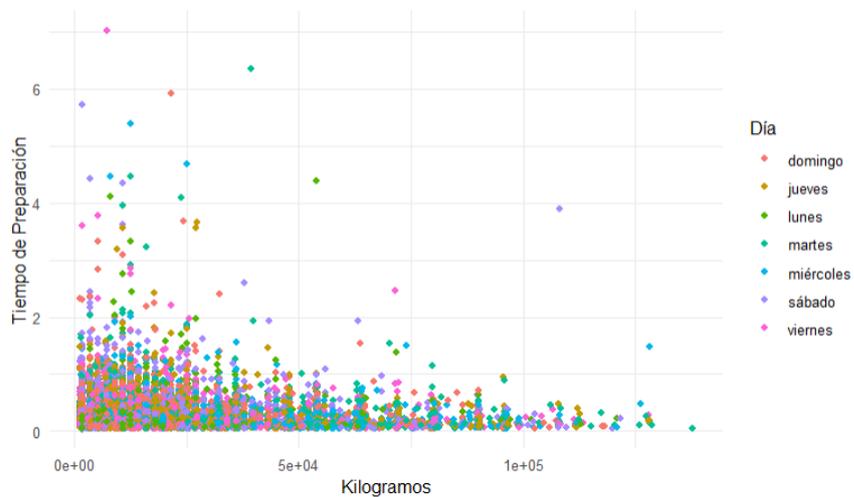


Ilustración 7: 53: Gráfico de Dispersión Kilogramos vs Tiempo de Preparación vs Día.

Rapidez vs Tiempo de preparación vs Día Ilustración 7: 54

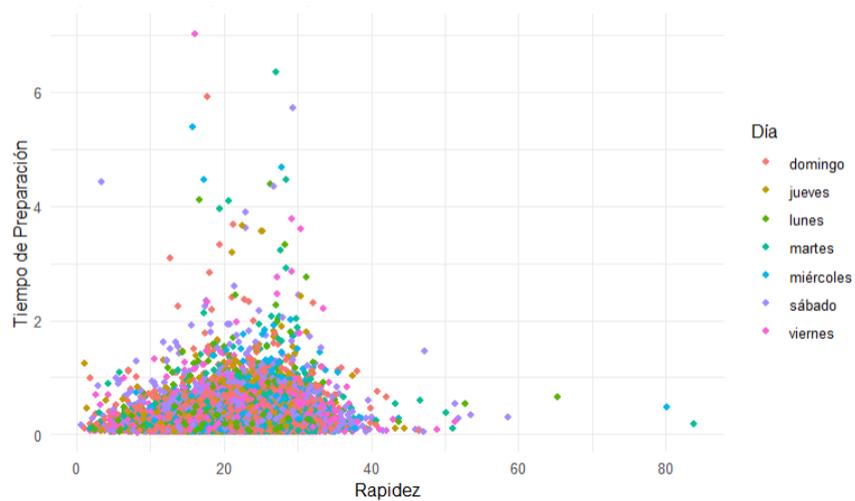


Ilustración 7: 54: Gráfico de Dispersión Rapidez vs Tiempo de Preparación vs Día.

Tiempo de Producción vs Kilogramos vs Mes Ilustración 7: 55

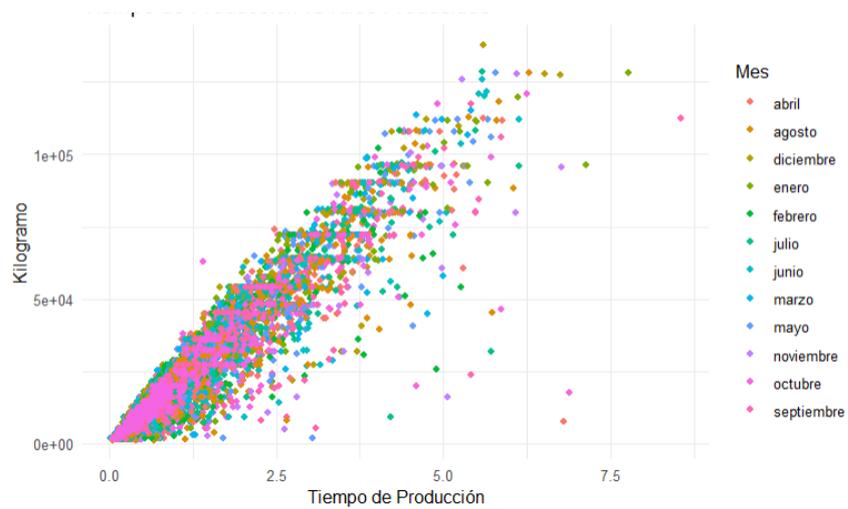


Ilustración 7: 55: Gráfico de Dispersión Tiempo de Producción vs Kilogramo vs Mes.

Tiempo de Producción vs Rapidez vs Mes Ilustración 7: 56

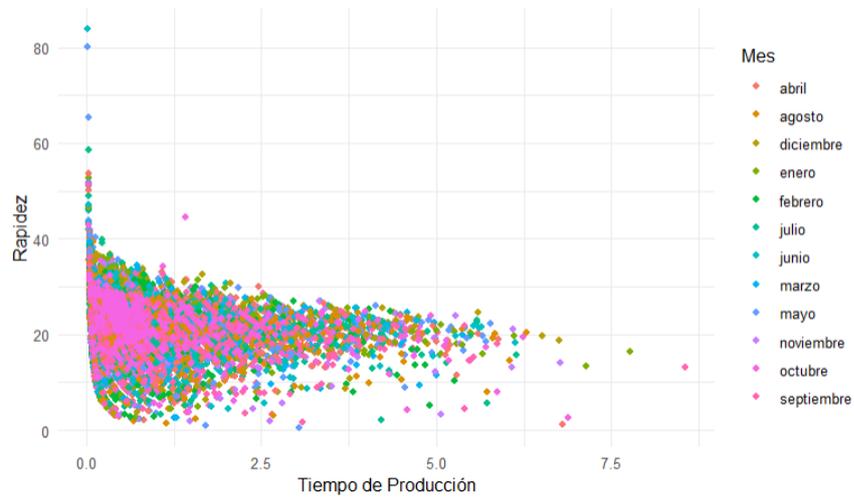


Ilustración 7: 56: Gráfico de Dispersión Tiempo de Producción vs Rapidez vs Mes.

Tiempo de Producción vs Tiempo de Preparación vs Mes Ilustración 7: 57

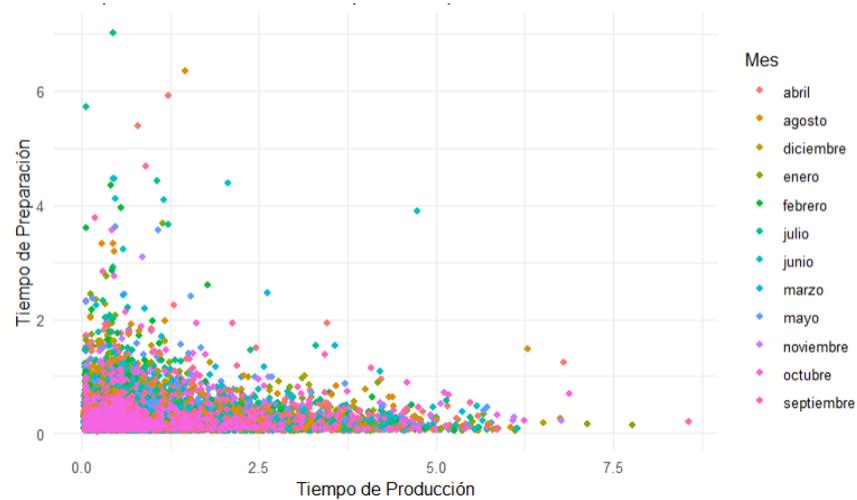


Ilustración 7: 57: Gráfico de Dispersión Tiempo de Producción vs Tiempo de Preparación vs Mes.

Kilogramos vs Rapidez vs Mes Ilustración 7: 58

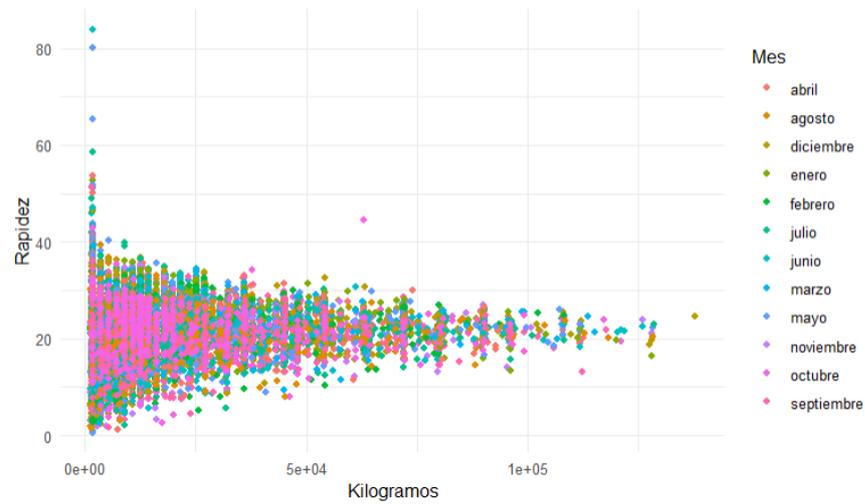


Ilustración 7: 58: Gráfico de Dispersión Kilogramos vs Rapidez vs Mes.

Kilogramos vs Tiempo de Preparación vs Mes Ilustración 7: 59

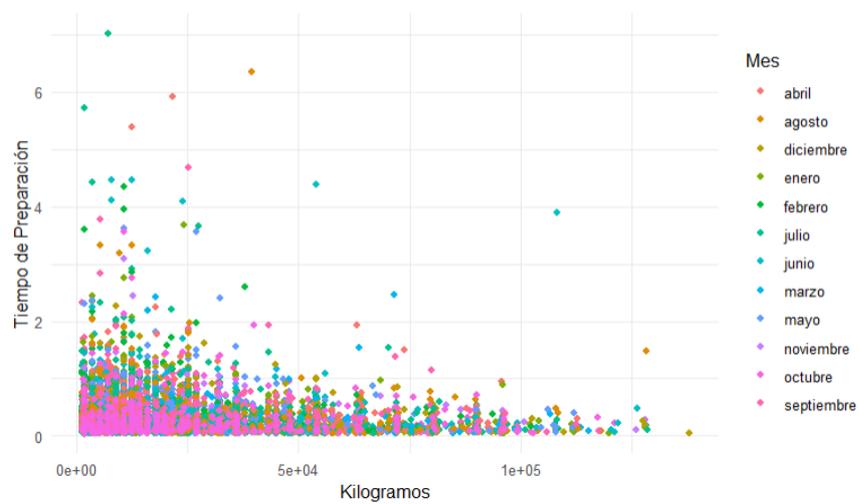


Ilustración 7: 59: Gráfico de Dispersión Kilogramos vs Tiempo de Preparación vs Mes.

Rapidez vs Tiempo de preparación vs Mes Ilustración 7: 60

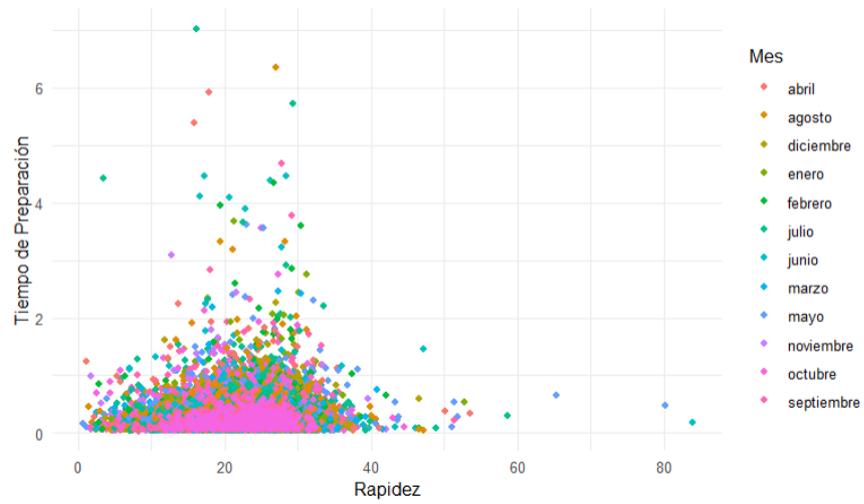


Ilustración 7: 60: Gráfico de Dispersión Rapidez vs Tiempo de Preparación vs Mes.

Tiempo de Producción vs Kilogramos vs Año Ilustración 7: 61

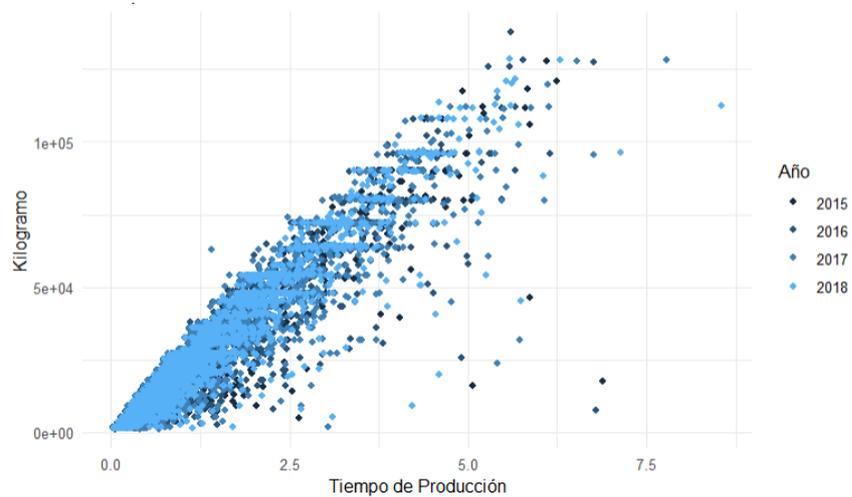


Ilustración 7: 61: Gráfico de Dispersión Tiempo de Producción vs Kilogramo vs Año.

Tiempo de Producción vs Rapidez vs Año Ilustración 7: 62

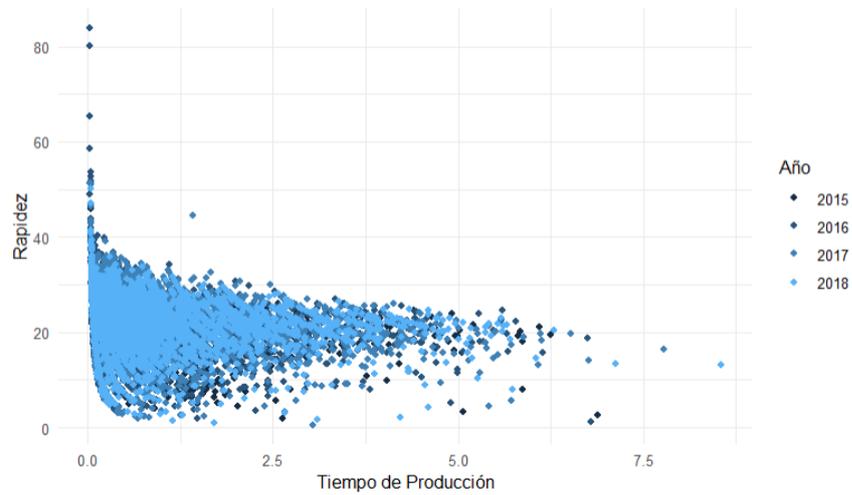


Ilustración 7: 62: Gráfico de Dispersión Tiempo de Producción vs Rapidez vs Año.

Tiempo de Producción vs Tiempo de Preparación vs Año Ilustración 7: 63

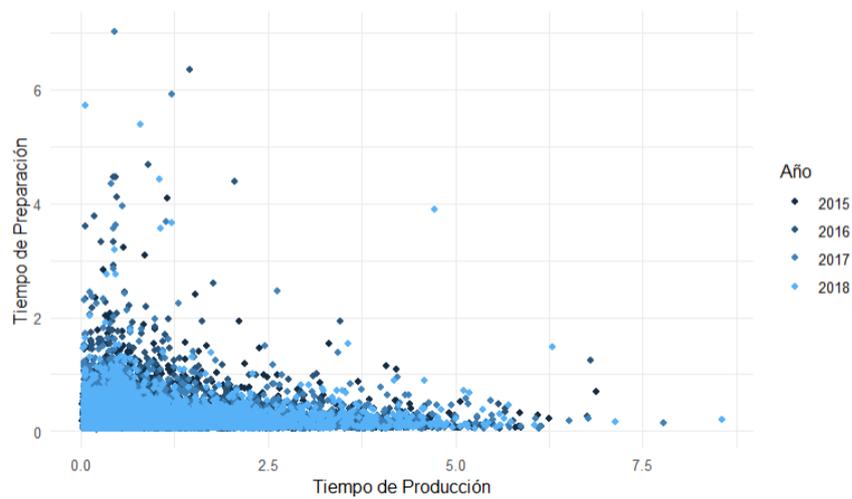


Ilustración 7: 63: Gráfico de Dispersión Tiempo de Producción vs Tiempo de Preparación vs Año.

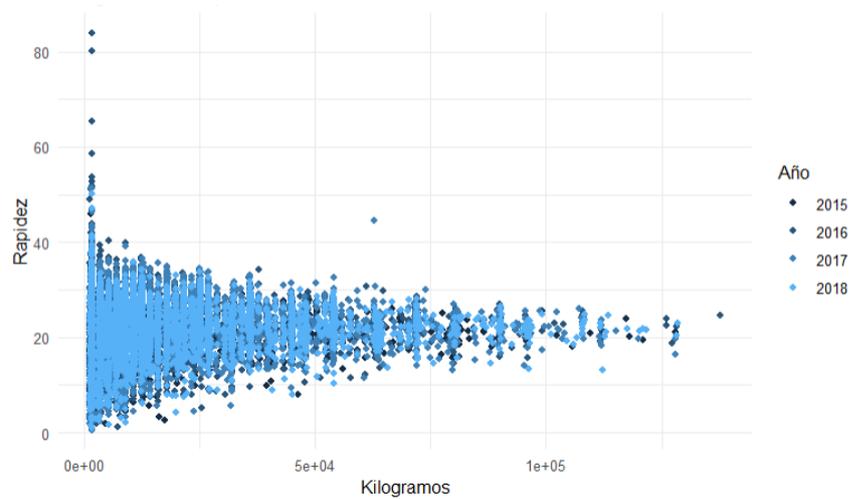
Kilogramos vs Rapidez vs Año Ilustración 7: 64

Ilustración 7: 64: Gráfico de Dispersión Kilogramo vs Rapidez vs Año.

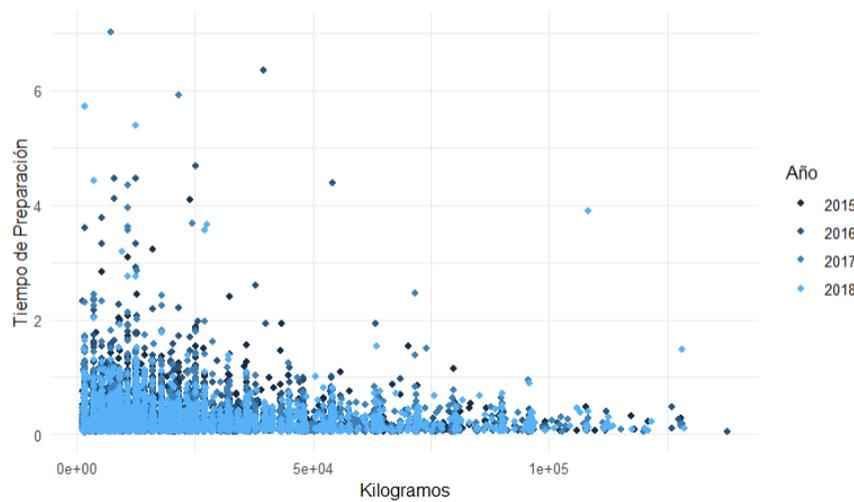
Kilogramos vs Tiempo de Preparación vs Año Ilustración 7: 65

Ilustración 7: 65: Gráfico de Dispersión Kilogramo vs Tiempo de Preparación vs Año.

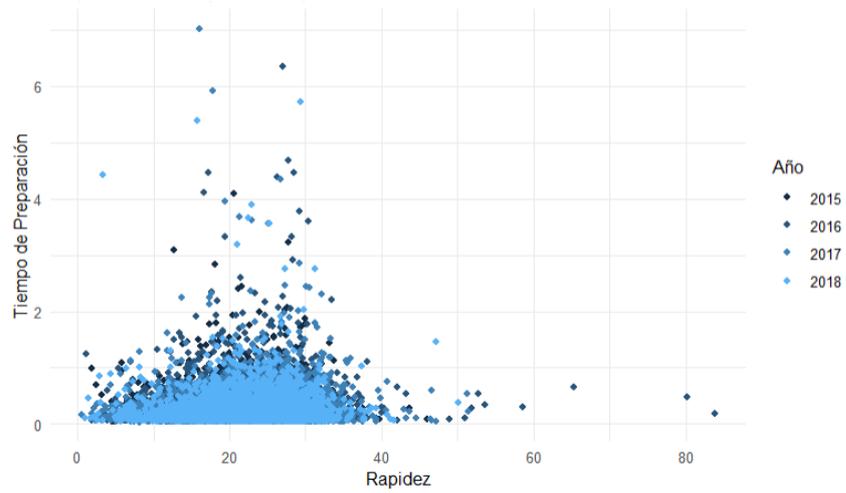
Rapidez vs Tiempo de preparación vs Año Ilustración 7: 66

Ilustración 7: 66: Gráfico de Dispersión Rapidez vs Tiempo de Preparación vs Año

7.5. Apéndice E: Matrices de Confusión Máquina Vector Soporte Kernel Radial Ilustración 7: 67 Ilustración 7: 68 Ilustración 7: 69 Ilustración 7: 70

```

Confusion Matrix and Statistics

      Reference
Prediction  -1   1
      -1  1898  1898
       1

      Accuracy : 0.5
      95% CI   : (0.484, 0.516)
      No Information Rate : 0.5
      P-Value [Acc > NIR] : 0.5065

      Kappa : 0

      Mcnemar's Test P-Value : <2e-16

      Sensitivity : 1.0
      Specificity : 0.0
      Pos Pred Value : 0.5
      Neg Pred Value : NaN
      Prevalence : 0.5
      Detection Rate : 0.5
      Detection Prevalence : 1.0
      Balanced Accuracy : 0.5

      'Positive' Class : 1

```

Ilustración 7: 67: Matriz de Confusión Kernel Radial Partición Entrenamiento Conjunto 1

```

Confusion Matrix and Statistics

      Reference
Prediction  -1   1
      -1 1479  145
       1  419 1753

      Accuracy : 0.8514
      95% CI   : (0.8397, 0.8626)
      No Information Rate : 0.5
      P-Value [Acc > NIR] : < 2.2e-16

      Kappa : 0.7028

      Mcnemar's Test P-Value : < 2.2e-16

      Sensitivity : 0.9236
      Specificity : 0.7792
      Pos Pred Value : 0.8071
      Neg Pred Value : 0.9107
      Prevalence : 0.5000
      Detection Rate : 0.4618
      Detection Prevalence : 0.5722
      Balanced Accuracy : 0.8514

      'Positive' Class : 1

```

Ilustración 7: 68: Matriz de Confusión Kernel Radial Partición Entrenamiento Conjunto 2

```

Confusion Matrix and Statistics

      Reference
Prediction -1  1
      -1  0  0
       1 474 474

      Accuracy : 0.5
      95% CI : (0.4677, 0.5323)
      No Information Rate : 0.5
      P-Value [Acc > NIR] : 0.513

      Kappa : 0

      Mcnemar's Test P-Value : <2e-16

      Sensitivity : 1.0
      Specificity : 0.0
      Pos Pred Value : 0.5
      Neg Pred Value : NaN
      Prevalence : 0.5
      Detection Rate : 0.5
      Detection Prevalence : 1.0
      Balanced Accuracy : 0.5

      'Positive' Class : 1

```

Ilustración 7: 69: Matriz de Confusión Kernel Radial Partición Prueba Conjunto 1

```

Confusion Matrix and Statistics

      Reference
Prediction -1  1
      -1 373  26
       1 101 448

      Accuracy : 0.866
      95% CI : (0.8427, 0.8871)
      No Information Rate : 0.5
      P-Value [Acc > NIR] : < 2.2e-16

      Kappa : 0.7321

      Mcnemar's Test P-Value : 5.153e-11

      Sensitivity : 0.9451
      Specificity : 0.7869
      Pos Pred Value : 0.8160
      Neg Pred Value : 0.9348
      Prevalence : 0.5000
      Detection Rate : 0.4726
      Detection Prevalence : 0.5791
      Balanced Accuracy : 0.8660

      'Positive' Class : 1

```

Ilustración 7: 70: Matriz de Confusión Kernel Radial Partición Prueba Conjunto 2

7.6. Apéndice F: Matrices de Confusión Máquina Vector Soporte Kernel Polinomial Ilustración 7: 71 Ilustración 7: 72 Ilustración 7: 73 Ilustración 7: 74

```

Confusion Matrix and Statistics

      Reference
Prediction  -1   1
      -1 1719 216
       1  179 1682

      Accuracy : 0.8959
      95% CI : (0.8858, 0.9055)
      No Information Rate : 0.5
      P-Value [Acc > NIR] : < 2e-16

      Kappa : 0.7919

      Mcnemar's Test P-Value : 0.07009

      Sensitivity : 0.8862
      Specificity : 0.9057
      Pos Pred Value : 0.9038
      Neg Pred Value : 0.8884
      Prevalence : 0.5000
      Detection Rate : 0.4431
      Detection Prevalence : 0.4903
      Balanced Accuracy : 0.8959

      'Positive' Class : 1

```

Ilustración 7: 71: Matriz de Confusión Kernel Polinomial Partición Entrenamiento Conjunto 1

```

Confusion Matrix and Statistics

      Reference
Prediction  -1   1
      -1 1479 124
       1  419 1774

      Accuracy : 0.857
      95% CI : (0.8454, 0.8679)
      No Information Rate : 0.5
      P-Value [Acc > NIR] : < 2.2e-16

      Kappa : 0.7139

      Mcnemar's Test P-Value : < 2.2e-16

      Sensitivity : 0.9347
      Specificity : 0.7792
      Pos Pred Value : 0.8089
      Neg Pred Value : 0.9226
      Prevalence : 0.5000
      Detection Rate : 0.4673
      Detection Prevalence : 0.5777
      Balanced Accuracy : 0.8570

      'Positive' Class : 1

```

Ilustración 7: 72: Matriz de Confusión Kernel Polinomial Partición Entrenamiento Conjunto 2

```

Confusion Matrix and Statistics

      Reference
Prediction -1  1
-1  429  47
 1   45 427

      Accuracy : 0.903
      95% CI : (0.8823, 0.921)
      No Information Rate : 0.5
      P-Value [Acc > NIR] : <2e-16

      Kappa : 0.8059

      McNemar's Test P-Value : 0.917

      Sensitivity : 0.9008
      Specificity : 0.9051
      Pos Pred Value : 0.9047
      Neg Pred Value : 0.9013
      Prevalence : 0.5000
      Detection Rate : 0.4504
      Detection Prevalence : 0.4979
      Balanced Accuracy : 0.9030

      'Positive' Class : 1

```

Ilustración 7: 73: Matriz de Confusión Kernel Polinomial Partición Prueba Conjunto 1

```

Confusion Matrix and Statistics

      Reference
Prediction -1  1
-1  361  28
 1  113 446

      Accuracy : 0.8513
      95% CI : (0.827, 0.8733)
      No Information Rate : 0.5
      P-Value [Acc > NIR] : < 2.2e-16

      Kappa : 0.7025

      McNemar's Test P-Value : 1.504e-12

      Sensitivity : 0.9409
      Specificity : 0.7616
      Pos Pred Value : 0.7979
      Neg Pred Value : 0.9280
      Prevalence : 0.5000
      Detection Rate : 0.4705
      Detection Prevalence : 0.5897
      Balanced Accuracy : 0.8513

      'Positive' Class : 1

```

Ilustración 7: 74: Matriz de Confusión Kernel Polinomial Partición Prueba Conjunto 2

7.7. Apéndice G: Matrices de Confusión Árbol de Clasificación Ilustración 7: 75 Ilustración 7: 76 Ilustración 7: 77 Ilustración 7: 78

```

Confusion Matrix and Statistics

      Reference
Prediction -1  1
-1  1413  406
 1   485 1492

      Accuracy : 0.7653
      95% CI   : (0.7515, 0.7787)
No Information Rate : 0.5
P-Value [Acc > NIR] : < 2.2e-16

      Kappa   : 0.5306

McNemar's Test P-Value : 0.008973

      Sensitivity : 0.7861
      Specificity : 0.7445
      Pos Pred Value : 0.7547
      Neg Pred Value : 0.7768
      Prevalence : 0.5000
      Detection Rate : 0.3930
      Detection Prevalence : 0.5208
      Balanced Accuracy : 0.7653

      'Positive' Class : 1

```

Ilustración 7: 75: Matriz de Confusión Árbol de Clasificación Partición Entrenamiento Conjunto 1

```

Confusion Matrix and Statistics

      Reference
Prediction -1  1
-1  1413  406
 1   485 1492

      Accuracy : 0.7653
      95% CI   : (0.7515, 0.7787)
No Information Rate : 0.5
P-Value [Acc > NIR] : < 2.2e-16

      Kappa   : 0.5306

McNemar's Test P-Value : 0.008973

      Sensitivity : 0.7861
      Specificity : 0.7445
      Pos Pred Value : 0.7547
      Neg Pred Value : 0.7768
      Prevalence : 0.5000
      Detection Rate : 0.3930
      Detection Prevalence : 0.5208
      Balanced Accuracy : 0.7653

      'Positive' Class : 1

```

Ilustración 7: 76: Matriz de Confusión Árbol de Clasificación Partición Entrenamiento Conjunto 2

```

Confusion Matrix and Statistics

      Reference
Prediction -1  1
      -1 347 106
       1 127 368

      Accuracy : 0.7542
      95% CI : (0.7255, 0.7813)
      No Information Rate : 0.5
      P-Value [Acc > NIR] : <2e-16

      Kappa : 0.5084

      McNemar's Test P-Value : 0.1901

      Sensitivity : 0.7764
      Specificity : 0.7321
      Pos Pred Value : 0.7434
      Neg Pred Value : 0.7660
      Prevalence : 0.5000
      Detection Rate : 0.3882
      Detection Prevalence : 0.5222
      Balanced Accuracy : 0.7542

      'Positive' Class : 1

```

Ilustración 7: 77: Matriz de Confusión Árbol de Clasificación Partición Prueba Conjunto 1

```

Confusion Matrix and Statistics

      Reference
Prediction -1  1
      -1 347 106
       1 127 368

      Accuracy : 0.7542
      95% CI : (0.7255, 0.7813)
      No Information Rate : 0.5
      P-Value [Acc > NIR] : <2e-16

      Kappa : 0.5084

      McNemar's Test P-Value : 0.1901

      Sensitivity : 0.7764
      Specificity : 0.7321
      Pos Pred Value : 0.7434
      Neg Pred Value : 0.7660
      Prevalence : 0.5000
      Detection Rate : 0.3882
      Detection Prevalence : 0.5222
      Balanced Accuracy : 0.7542

      'Positive' Class : 1

```

Ilustración 7: 78: Matriz de Confusión Árbol de Clasificación Partición Prueba Conjunto 2

