

Tabla de Contenido

1. Introducción	1
1.1. Timejobs	1
1.1.1. Contexto	1
1.1.2. Tipos de jobbers	2
1.1.3. Plataformas Digitales	3
1.1.4. Reclutamiento	3
1.1.5. Problemática Actual en Reclutamiento	4
1.2. Objetivos	5
1.2.1. Objetivo General	5
1.2.2. Objetivos Específicos:	6
1.3. Estructura del Documento	6
2. Estado Del Arte	7
2.1. On Boarding	7
2.2. Big Data	8
2.3. Inteligencia de Negocios	8
2.4. Data Analítica	9
3. Marco Teórico	11
3.1. Ciencia de Datos e Inteligencia Artificial	11
3.1.1. Minería de Datos	11
3.1.2. Aprendizaje de Maquina	12
3.1.3. Tipos de Aprendizaje	13
3.2. Algoritmos de Aprendizaje de Maquina	13
3.2.1. Red Neuronal	14
3.2.1.1. Perceptrón	14
3.2.1.2. Perceptrón Multicapa	17
3.2.2. Árboles y Bosques	19
3.2.2.1. Árboles de Decisión	19
3.2.2.2. Random Forest	21
3.2.2.3. AdaBoost	22
3.2.2.4. Gradient Boosting	23
3.2.3. Naives Bayes	24
3.2.3.1. Clasificador Bayesiano	24
3.2.4. Métricas de Desempeño	25
3.3. Metodologías para proyectos de Ciencia de Datos	26
3.3.1. KDD	26

3.3.2. CRISP-DM	28
4. Metodología	30
4.1. Caracterización	30
4.2. Modelos	31
4.3. Implementación Modelo	32
5. Procesamiento y Modelamiento	33
5.1. Caracterización del usuario	33
5.2. Interacción con la plataforma	34
5.3. Distribución de inducciones	36
5.4. Modelos	37
6. Resultados y Análisis	38
6.1. Resultados	38
6.1.1. Redes Neuronales	38
6.1.2. Arboles y Bosques	41
6.1.3. Naives Bayes	45
6.1.4. Resultados completos modelos	46
6.2. Análisis	48
6.2.1. Redes Neuronales	48
6.2.2. Arboles de Decisión	48
6.2.3. Random Forest y Gradient Boosting	49
6.2.4. Naives Bayes	50
6.2.5. Desempeños Generales	50
7. Implementación	52
8. Conclusiones	55
Bibliografía	56
Anexos	58
A. Distribución variables de entrada	58
B. Caracterización	59
C. Agrupamiento de Usuarios que no terminan el registro	61
D. Resultados Modelos con bajo desempeño	62