

Inteligencia Artificial

Proyecto Final **Predicción Laboral según la** **hoja de Vida**

Gabriela Garcés Díaz

—

Profesor:

Francisco Carlos Calderon Bocanegra

—

Noviembre 2022

Introducción

Buscar y obtener trabajo es una tarea compleja para cualquier persona, es por ello que el objetivo de este proyecto es estimar con facilidad a partir de la experiencia necesaria, competencias y habilidades claves el cargo que se puede esperar ocupar considerando un dataset sobre los títulos de los puestos, los requisitos de experiencia laboral, las competencias clave, las categorías de funciones, las ubicaciones, las áreas funcionales, los sectores, las funciones, las longitudes, las latitudes y los salarios, y así determinar los trabajos perfectos para cada persona.

Desarrollo de la Implementación

Cabe resaltar que este problema es de tipo supervisado y se solucionó por medio de máquinas de soporte vectorial, que es un método de clasificación.

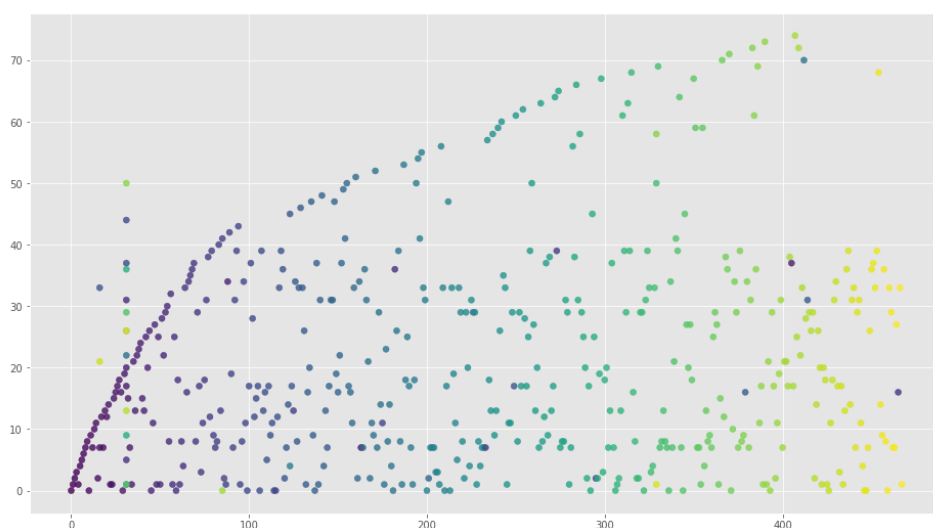
Para la manipulación del dataset, en primer lugar, se realizó la limpieza donde se eliminó los valores faltantes o NaN de existir, y de hecho se eliminó la segunda columna que tenía valores sin utilidad, para ello, se revisó varios estadísticos de la distribución de los datos. El dataset nos proporciona 500 muestras con 11 características; por lo que adicionalmente se calculó el conteo de cada característica.

Una vez analizados los datos, se discretizaron, es decir se pasó los valores categóricos a numéricos. Este archivo se exportó y se guardó en el disco interno bajo el nombre “discrete_jobs”.

En cuanto al desarrollo del modelo, se tiene los datos limpios tales que por columnas tenemos nuestras características y por filas cada muestra las etiqueta(s) “Y”; considerando la gran cantidad de características y nuestro objetivo de predecir el cargo del trabajo, se diseñó y se X de la siguiente manera

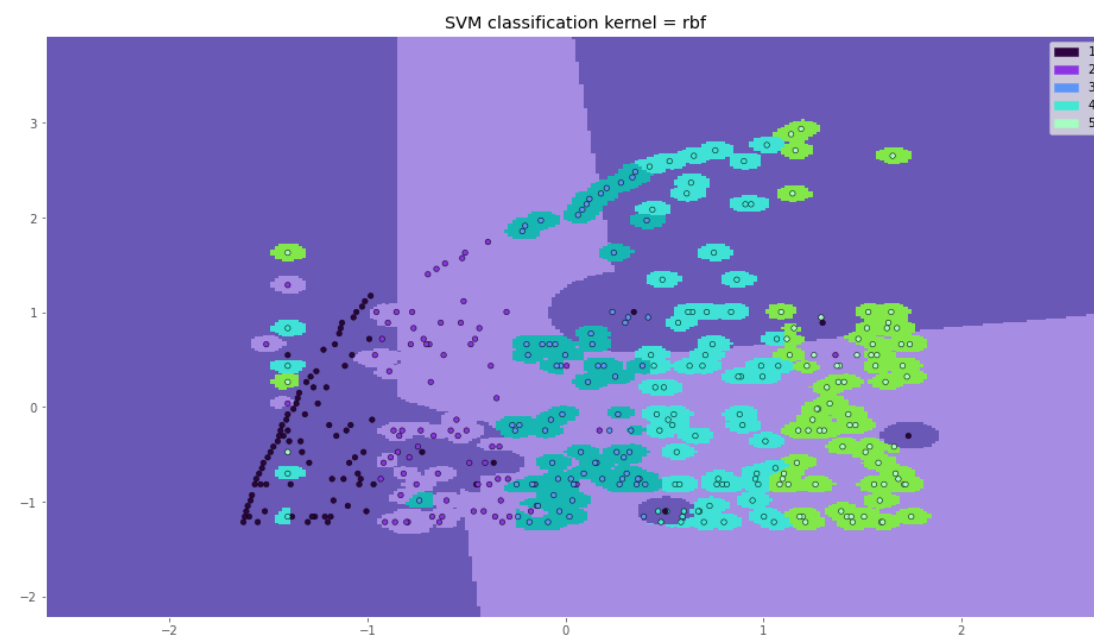
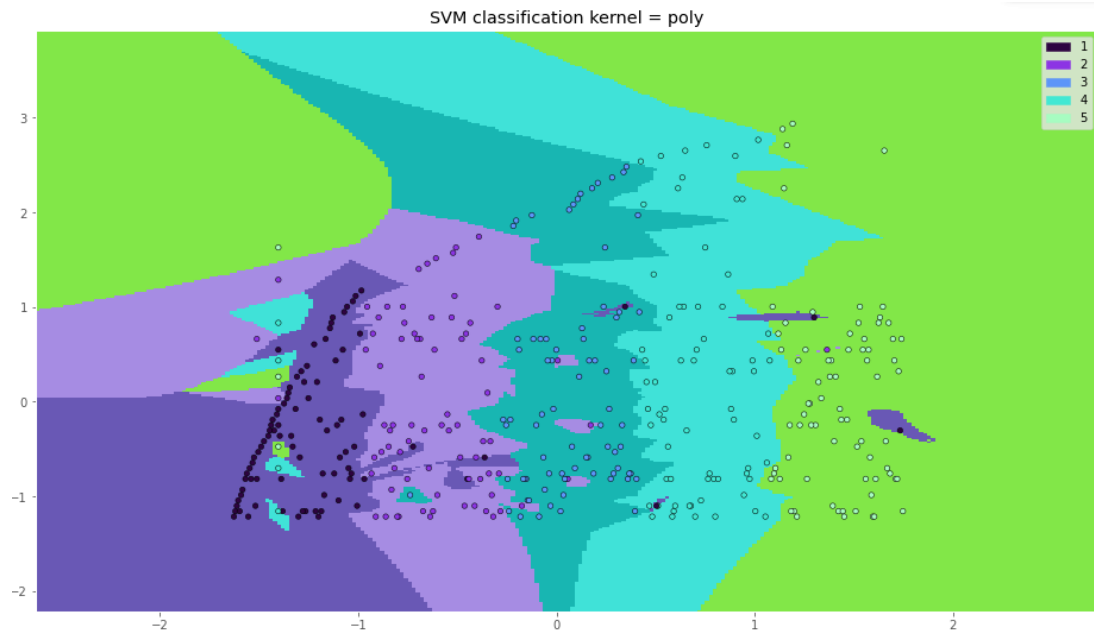
```
X = dataframe[['Key Skills', 'Job Experience Required']].values  
y = dataframe['Job Title'].values
```

Posteriormente, se usó el 20% de los datos para el entrenamiento y se graficaron los datos organizados.



Luego, se entrenó el modelo usando el kernels= 1 y kernels= 2 (poly y rbf) para determinar la mejor predicción.

Resultados



Se evidencia que el mejor modelo fue usando poly ya que para cada uno de los cinco conjuntos existe zonas de acuerdo con su distribución, en comparación de cuando se utilizó rbf que parece que se sobreentrenara el modelo ya que las zonas para tres de los conjuntos (3, 4 y el 5) son prácticamente la posición de cada una de muestras.

Conclusiones

Dada la magnitud y la cantidad de este dataset se puede lograr realizar numeros y complejos modelos para la predicción de un cargo laboral, o inclusive predicciones sobre la localización o el salario estimado de cierto cargo. Asi mismo, fue indispensable la manipulación correcta de los datos para asi obtener un modelo coherente el cual logre predecir el cargo laboral que se puede ejercer.