

Aplicación de Modelos de Aprendizaje Supervisado en la Industria de Confección

Julian Camilo Restrepo Angel

4 de julio de 2025

Resumen

Este proyecto aplica técnicas de ciencia de datos y modelos de aprendizaje supervisado para predecir y clasificar la productividad de equipos en una planta de confección. Se utilizaron modelos como regresión lineal, regresión logística, K-Vecinos más cercanos (K-NN) y árboles de decisión, evaluando su desempeño y aplicabilidad al contexto industrial.

1. Acción 1: Regresión Lineal

- Se dividió la base de datos en 80 % entrenamiento y 20 % prueba.
- Se aplicó un modelo de regresión lineal para predecir la productividad real con base en el número de trabajadores y el tiempo extra en minutos.
- Se realizaron predicciones sobre la base de prueba.
- Se generó un diagrama de dispersión con valores predichos vs. valores reales de productividad.
- Se evaluó el modelo utilizando el Error Cuadrático Medio (ECM) y la Raíz del Error Cuadrático Medio (RMSE).

2. Acción 2: Regresión Logística

- Se creó la variable **nivel_productividad**:
 - Alto si productividad real > 0.7
 - Bajo si productividad real ≤ 0.7
- Se entrenó un modelo de regresión logística con las variables: número de trabajadores y tiempo extra.
- Se realizaron predicciones sobre la base de prueba.
- Se calculó la precisión (accuracy) del modelo.

3. Acción 3: K-Vecinos más Cercanos (K-NN)

- Se aplicó el modelo KNN con la base de entrenamiento para predecir **nivel_productividad**.
- Se utilizaron como variables independientes: número de trabajadores y tiempo extra.
- Se evaluó el modelo sobre la base de prueba mediante precisión (accuracy).

4. Acción 4: Árboles de Decisión (CART)

- Se entrenó un árbol de decisión usando todas las características para predecir **nivel_productividad**.
- Se visualizó el árbol utilizando `graphviz()`.
- Se realizaron predicciones sobre la base de prueba.
- Se calculó la precisión (accuracy) del modelo.

5. Conclusiones

- Se identificó una relación significativa entre el número de trabajadores, tiempo extra y la productividad.
- Los modelos de clasificación permiten tomar decisiones anticipadas sobre el rendimiento esperado.
- Herramientas como los árboles de decisión permiten una visualización clara para apoyar decisiones operativas.