

UNIVERSIDAD DEL VALLE DE GUATEMALA

Minería de datos



Laboratorio #8

Redes neuronales

Bryan Carlos Roberto España Machorro - 21550

Javier Alejandro Prado Ramirez - 21486

Gabriel García

Comparación de resultado con regresión con redes neuronales

El primer modelo de la red neuronal tuvo una arquitectura de capas densamente conectadas, con un optimizador RMSprop, teniendo como pérdida el cuadrado medio (MSE) y como métrica de la evaluación se utilizó el error absoluto medio (MAE)

El segundo modelo tuvo una arquitectura de capas densamente conectadas con regularización Dropout, optimizador adam y al igual que el anterior utilizó MSE para la función de pérdida y MAE para la métrica de evaluación.

- MAE
 - Primer Modelo: Entrenamiento: ~57,840, Prueba: ~56,379
 - Segundo Modelo: Entrenamiento: ~57,509, Prueba: ~56,039
- MSE
 - Primer Modelo: ~6,179,734,960
 - Segundo Modelo: ~6,182,751,366
- R^2
 - Ambos tiene un R^2 muy cercano a cero indicando que los modelos no capturan la variabilidad de los datos de manera significativa

Efectividad:

Con respecto a la efectividad ambos modelos muestran resultados muy similares en rendimiento. Lamentablemente el rendimiento es insatisfactorio, ya que los valores mostrados anteriormente son bastante altos y el coeficiente de determinación R^2 está cercano a cero, lo que nos muestra que los modelos no logran explicar la variabilidad de los datos de manera adecuada.

Procesamiento:

El tiempo de procesamiento varía dependiendo del dispositivo, ya que este va de la mano con el hardware y el tamaño del conjunto de datos, pero en general los modelos pueden requerir bastante tiempo para el entrenamiento, tardando alrededor de 5 minutos.

Errores:

- Sobreajuste: Dado que los modelos no lograron generalizar bien a partir de los datos de entrenamiento a los datos de prueba es probable que hayamos tenido un sobreajuste
- Subajuste: Los modelos pueden estar infraentrenados debido a su estructura y la falta de información
- Errores en la predicción: Los errores más significativos ocurren en las predicciones de precios de venta de casas. Estos errores pueden ser por la complejidad del problema, la falta de características importantes en los datos, la presencia de valores atípicos en los datos.

Conclusiones:

Aunque los modelos de redes neuronales pueden ser útiles para las tareas de regresión como la predicción de precios de venta de casas, es importante tener en cuenta que su efectividad varía dependiendo de varios factores, incluida la calidad de los datos y la selección correcta.

Es importante realizar un análisis de los resultados y evaluar la efectividad de los modelos en función de métricas adecuadas, así como comprender las limitaciones y posibles sesgos en el proceso del modelado