

Metro Interstate Traffic Volume

MARIA RUBIO RUBIO.MARIA@GALILEO.EDU, FREDY MARROQUIN
17001575@GALILEO.EDU, CRISTINA VILLAGRAN 17005199@GALILEO.EDU

Problema

En la autopista interestatal 94 en dirección oeste para la estación 301 de DoT de MN podemos observar como el volumen de tráfico crece cada año. Nuestro objetivo es predecir el volumen de tráfico diario para los próximos años utilizando los siguientes parametros; Festividades, temperatura[K] promedio en kelvin, lluvia [mm/h]. Nieve[mm/h], porcentaje de nubes, clima, fecha y hora, volumen de trafico.

Modelos

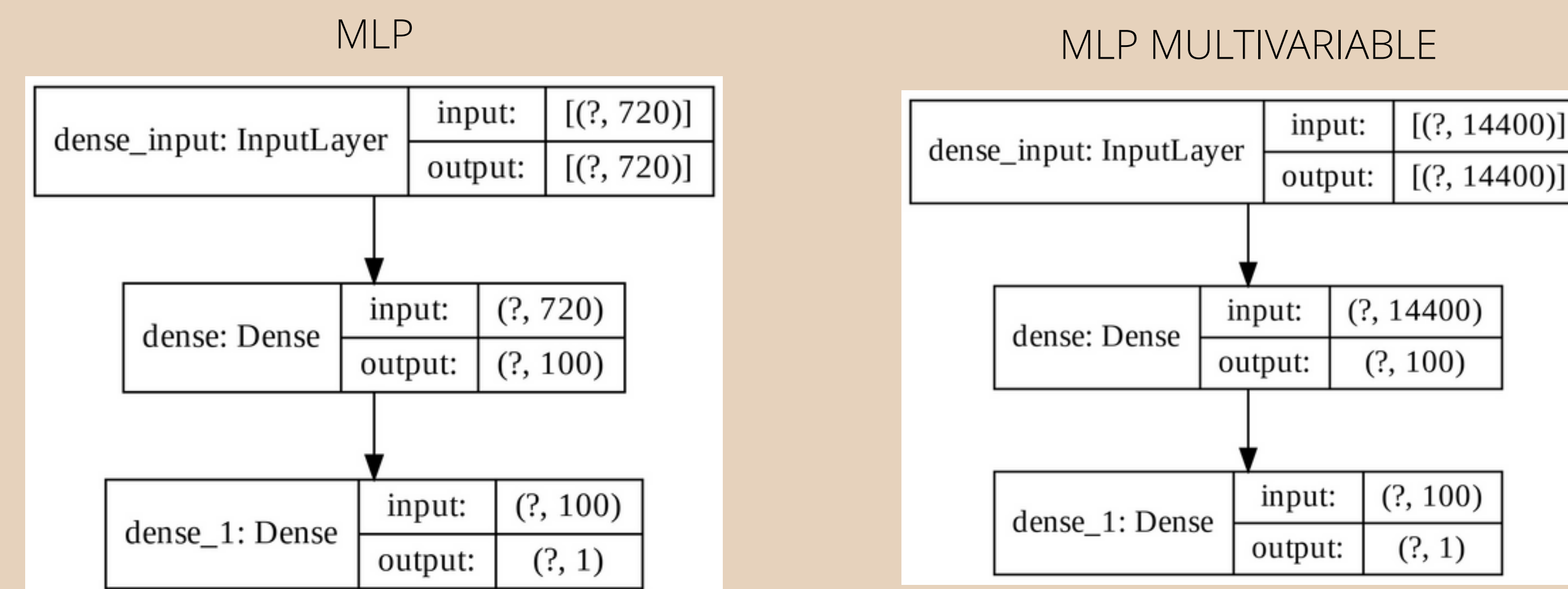
Para este problema decidimos usar varios modelos, para poder atacar el problema de distintas perspectivas y poder obtener mejores resultados, los modelos utilizados fueron.

- Random Forest
- Extra trees
- AdaBoost
- Multi-layer Perceptron univariable
- Multi-layer Perceptron multivariable

Analisis de datos

Decidimos separar los datos de una forma para los modelos de Random, Forest, Extra trees y Adaboost, de modo que despues de realizar nuestro respectivo análisis nos quedamos unicamente con los campos: holiday, temp, clouds_all, hour y traffic_volume. Para los MLP se distribuyeron de la misma manera pero progresivamente

Arquitecturas principales



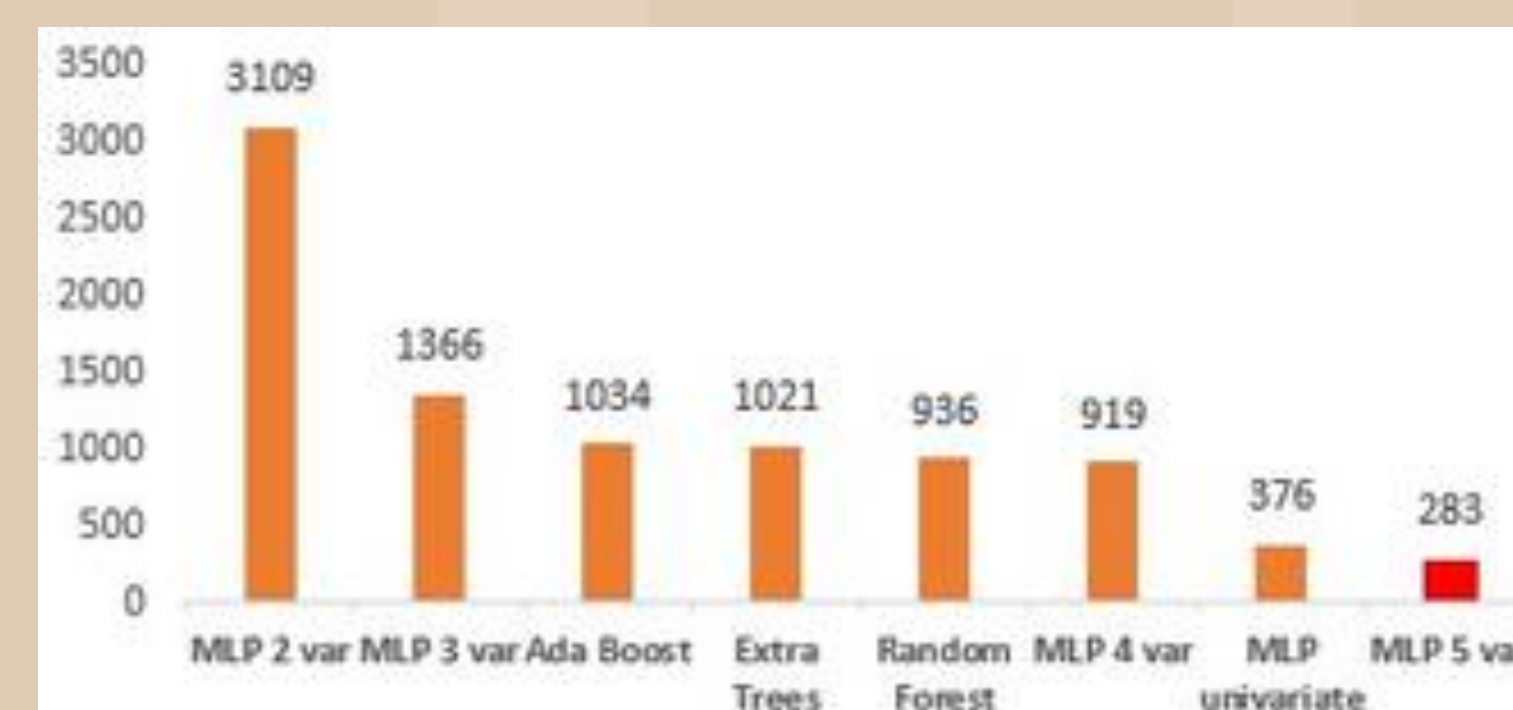
conclusiones

Un modelo MLP se adapta mucho mejor al problema ya que con base a las predicciones es el que más cerca esta de predecir con mayor precision cuanto volumen de trafico existira para el año siguiente, entre más variables significativas se agreguen el modelo ira mejorando.

Se demostró que es posible determinar el volumen de tráfico de la Interestatal 94 mediante el uso de las variables temperatura, volumen de tráfico, lluvia, nieve y día festivo, que nos indica la matriz de correlación

Resultados

Para poder obtener resultados, de 27,860 datos, usamos para entrenar 9306, para validar 10650 y para la prueba 7949. Se tomó esta información del año 2016 al 2018. Para poder aplicar los algoritmos que mencionamos. Ya aplicados los algoritmos y tener nuestros modelos, decidimos entrenar y medir con RMSE, los resultados obtenidos. Que fueron los siguientes.



Modelo	RMSE
MLP 2 var	3109
MLP 3 var	1366
Ada Boost	1034
Extra Trees	1021
Random Forest	936
MLP 4 var	919
MLP univariate	376
MLP 5 var	283

Investigación futura

Para investigaciones futuras tenemos 3 importantes postulados a destacar:

- Probar nuevos modelos como un LSTM
- Tratar de buscar un learning rate que se adapte mejor
- Aleatorizar el dataset o buscar maneras alternativas de mover la data para obtener mejores resultados

Referencia principal

Machine Learning Approaches for Traffic Volume Forecasting: A Case Study of the Moroccan Highway Network (November 21, 2017). <https://arxiv.org/>