

Estado del proyecto MobyTensorFlow

/Concepts

Finalizado

/DSL

Como te comente, lo realice todo en un notebook siguiendo un ejemplo del libro que nos pasaste "Hands-On Machine Learning with Scikit-Learn & TensorFlow" de Aurélien Géron. Es un poco complejo y algunas cosas no las entiendo, pero me sirvió para aprender lo básico de como se trabaja con las librerías habituales tipo pandas, numpy, etc.

Solo tendría que separarlo en los respectivos notebooks, que es lo que estoy haciendo ahora.

/Use Case

1. /Setup

Hay un archivo con instrucciones para configurar el entorno de forma que no de problemas. Estoy trabajando en Linux y no me ha dado ningún problema.

2. /Code

Esta todo el código que he usado.

3. /Problem

Describo que se trata de un problema de movilidad, los datos que uso, el lenguaje, etc.

4./DMP

El archivo UC_DMP es el importante, los demás son test que hacia.

En cuanto al proceso que he seguido ha sido el siguiente:

He añadido la columna objetivo. He tenido que reducir la precisión milimétrica que daban los datos ya que de otra forma la proporción de datos era exageradamente desproporcionada (99.99% no estaba en casa). Además tampoco tenía sentido que la casa estuviera localizada en un milímetro.

El problema del dato de las fechas que te comente lo he resuelto convirtiéndolas a números enteros.

Normalize los datos

En cuanto a la red neuronal, vi que el problema se trataba de una clasificación binaria, por lo que he escogido una función de activación sigmoidea para la neurona de salida y como función de pérdida la binary_cross_entropy. Te lo explico en el notebook, pero estoy seguro de que es lo correcto para este caso.

Para evaluar los resultados he usado la matriz de confusión y la curva ROC. Lo explico en el notebook

Hay explicaciones más detalladas de todo en el notebook.

5./Results

Los resultados están en /DPM, ya que necesitaba la salida de la red neuronal para poder visualizarlos