

Problema predictivo a resolver:

En esta competencia <https://www.kaggle.com/competitions/sberbank-russian-housing-market/overview> Sberbank nos desafía a desarrollar algoritmos que utilicen un amplio espectro de características para predecir los precios de los bienes raíces. Con un rico conjunto de datos que incluyen datos sobre vivienda y patrones macroeconómicos. Un modelo de pronóstico más preciso permitirá a Sberbank brindar más certeza a sus clientes en una economía incierta.

Dataset a utilizar:

El dataset cuenta con una amplia variedad de columnas que poseen muchos tipos de datos que van desde datos macro económicos (IPC, PIB . . .) hasta una colección de características sobre el vecindario circundante de cada propiedad y algunas características que son constantes en cada subárea. En el dataset la mayoría de los nombres de las funciones se explican por sí solas.

Métrica de desempeño:

Como métrica se va a utilizar un [RMSLE](#) entre sus precios previstos y los datos reales. La variable objetivo, denominada precio_doc en el conjunto de entrenamiento, es el precio de venta de cada propiedad.

Desempeño deseable:

El desempeño deseable a priori para el modelo es de un 80% dado que se cuenta con una amplia cantidad de datos, los cuales hacen posible un modelo robusto.