airbnb

As an estimated $38 billion company, according to Forbes in 2018, Airbnb has shocked the hospitality world.
With Airbnb’s presence in 192 worldwide countries and 81,000 cities, with just internet access, you can pretty much use the service anywhere.

Project Flow

Business Understanding:

Create a model that predicts the holiday destination for a new Airbnb User.

Data Understanding:

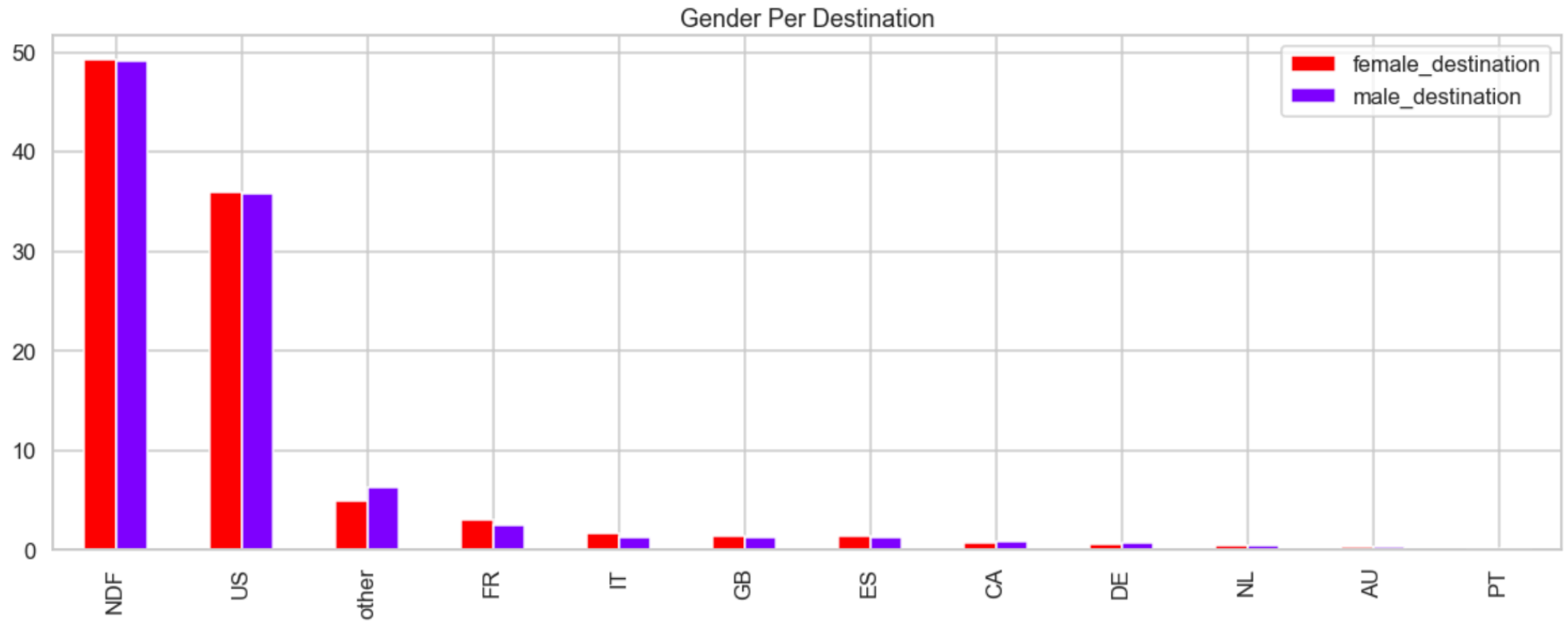
Clean the data, perform some feature engineering and EDA.

Modelling:

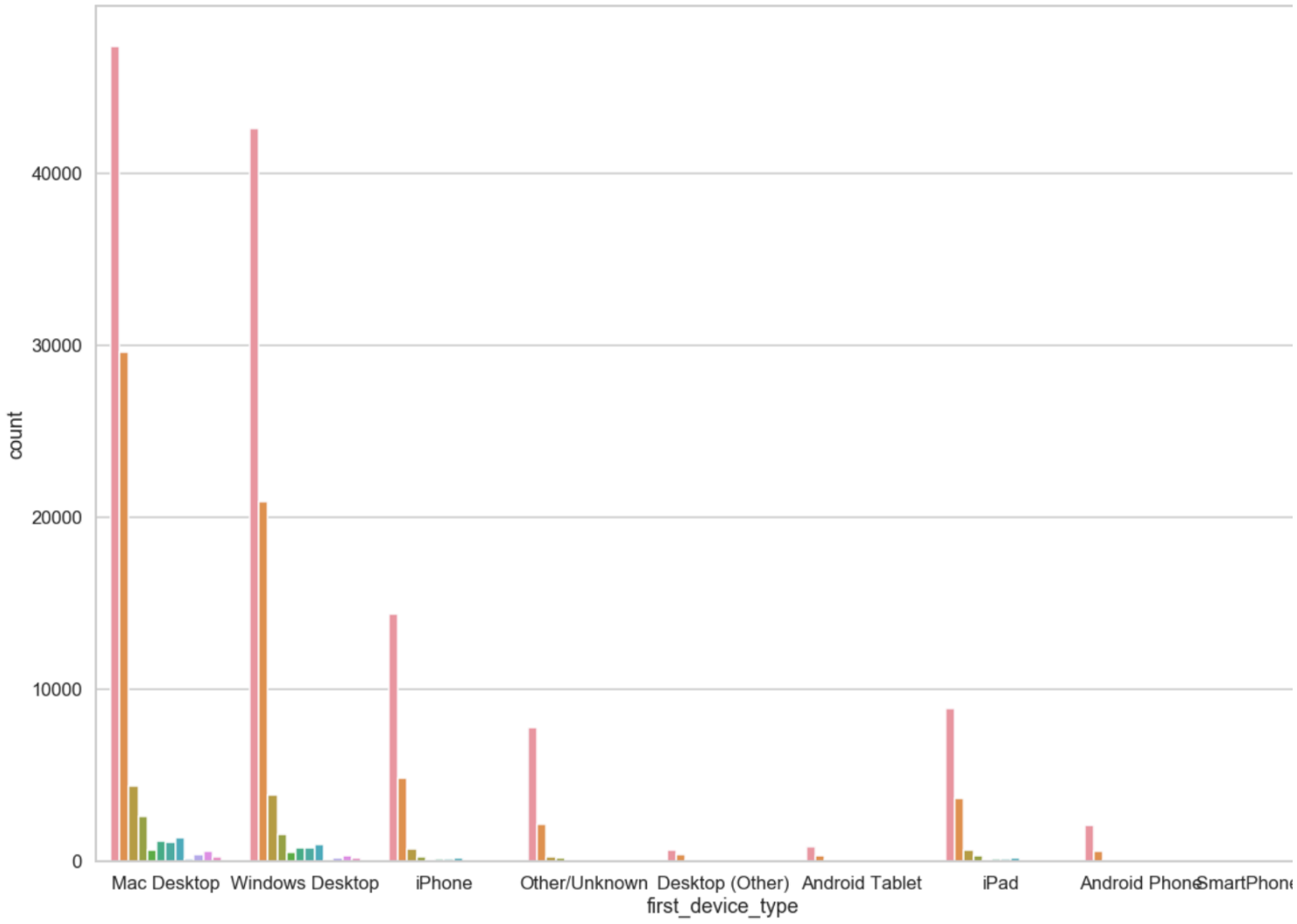
Logistic Regression, KNN, Decision Tree.

Evaluation:

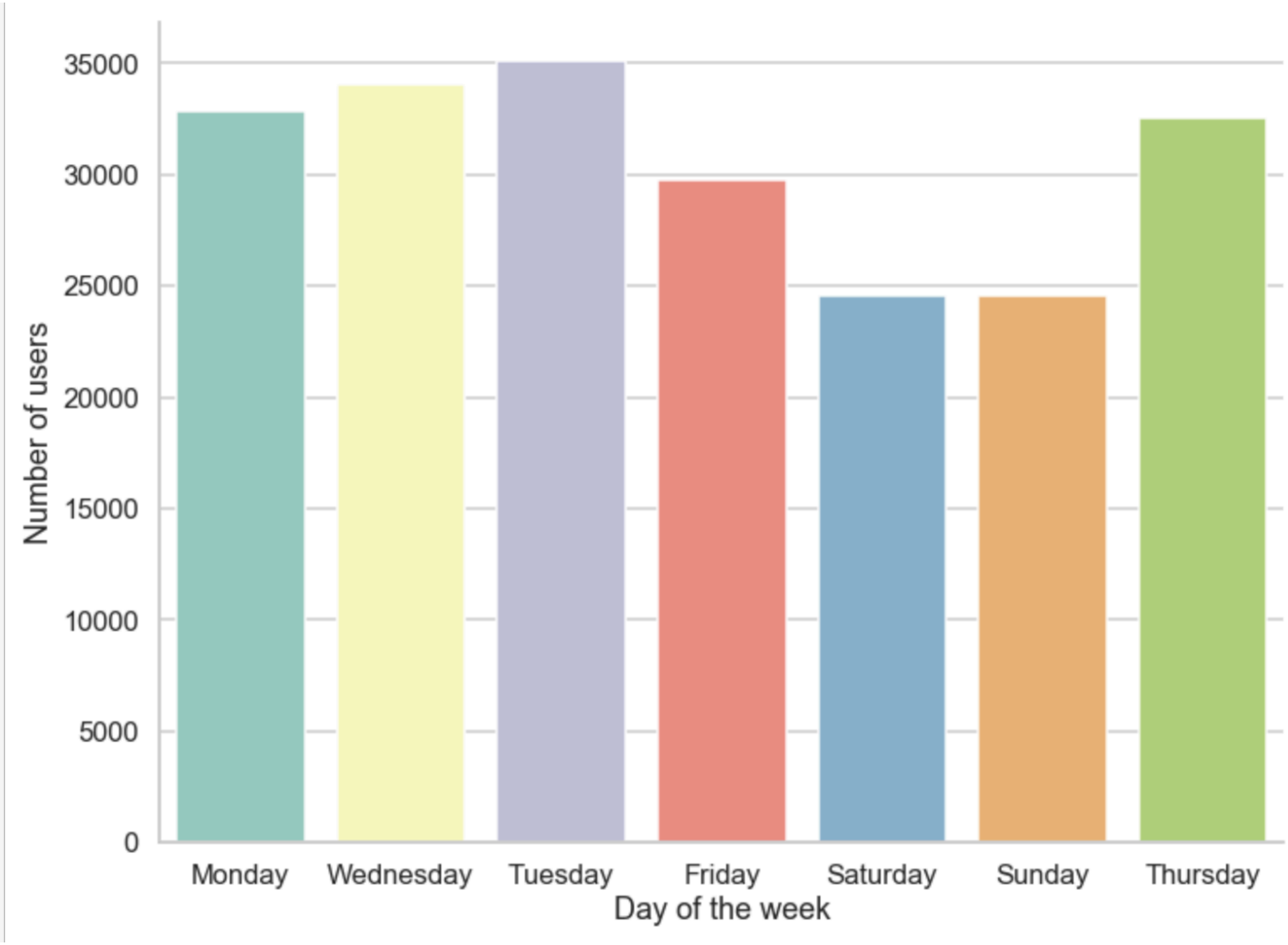
Choosing the best predictive model by analysing the evaluation metric Roc_Auc



Given 10 possible destinations, is it possible to predict where a new user will spend his next holidays?



Which are the most used devices when a new user approaches for the first time Airbnb?



Which are the days in the week with the major activity in terms of creating a new user profile on the Airbnb platform?

Evaluation

My winning model is Logistic Regression tuned with hyperparameter, giving a score of 0,804.

LOGISTIC REGRESSION

- **Basic model: 0,77 score of ROC_AUC**
- **Hyperparameter tuned model: 0,804 score of ROC_AUC**

KNN

- **Basic model: 0,682 score of ROC_AUC**
- **Hyperparameter tuned model: 0,77 score of ROC_AUC**

DECISION TREE

- **Basic model: 0,635 score of ROC_AUC**
- **Hyperparameter tuned model: 0,672 score of ROC_AUC**

Business insights from the winning model

Thanks to my model I am able to predict whether a new Airbnb user will go to the US or not.

The result of my winning model can be be improved. However with a score of the ROC_AUC of 0,565, Airbnb can classify better the customers and implement targeted marketing campaigns.

Future improvements

For the new users that are not going to the US, predict which country in particular will be their next destination.

Implement the additional model Random Forrest

Perform more EDA and gain from it other useful insights to increase the revenue of the company

Contact info

- soledadmusella@gmail.com
- linkedin.com/in/soledad-musella-bb8665104