Maximizing Wine Quality with Machine Learning

Douro Valley

REPÚBLICA PORTUGUESA

Barcelona Technology School

Current situation

Every year Douro Valley places millions of euros on wine harvesting with the hope that this investment will generate good crops and ultimately the best wine quality.

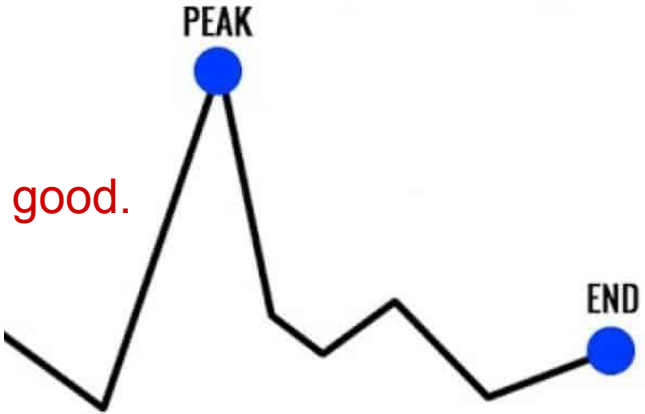
Better wine quality >> More enjoyment >> Better sales

Makes sense.

But, we know this doesn't happen in reality.

There are years where the wine is terrible or not so good.

Bad product >> Bad experience >> Low sales.



Think about all the investment you made…

All the time you kept the wine resting in barrels…

All the logistics and operations you put in place thinking that we would have certain amount of sales…

We are here to help :)

We predict the quality of your wine.

And you focus on selling.

By using a public dataset of wine metrics tracked by Douro Verde lab, we are able to predict the quality of white and red wines from North of Portugal.

Results?

With an accuracy around 80% we built a mode that is able to match the quality score of the wine specialists that labeled the wines dataset: low, medium, and high quality.

Why using our model?

Less unpredictability

Less costs

Less turnover

More preparation

More brand recognition

More money

Better Douro Valley

So how do we build our model?

Dataset: 6496 wine tests

11 Input variables (based on physicochemical tests): fixed acidity, volatile acidity, citric acid, residual sugar, chlorides, free sulfur dioxide, total sulfur dioxide, density, sulphates, alcohol, type of wine

1 Output variable (based on sensory data): quality (score between 0 and 10)

Model: Supervised learning with Random Forest Classifier using Standardization and SMOTE for oversampling.

Git repo: https://github.com/zemariatrindade/WineClassifier

How can you help us make our predictions better?

Our training dataset contained very few high quality wines, and for that reason the model wasn't able to capture the relationships across the features with the quality score.

As such, with a very small effort on getting high quality wine samples, we are confident to bring the accuracy above 90%.

Furthermore, we believe that are other features will make the predictions more accurate: type of barril, type of shipment, grape type, wine brand, etc.

help us helping you.
let’s work :)