

Projet recettes de bières

Sylvain Gault

13 septembre 2023

1 Introduction

Ce projet est à réaliser par groupe trois en Python en utilisant au moins scikit-learn et matplotlib. Si besoin, vous pourrez utiliser Pandas, et tensorflow. Il sera à rendre sur github et vous en ferez une présentation le Mardi 19/09/2023 dans l'après-midi.

2 Cahier des charges

2.1 Introduction



Vous êtes embauché par la société BREW (Beer Recipes Enthusiasts Workshop) afin d'accélérer le développement de nouvelles recettes de bières en utilisant l'IA intelligente et l'apprentissage Machine Learning par ordinateur. (Visiblement, le RH ne sait pas trop de quoi il parle.)

Votre mission si vous l'acceptez, sera de prédire l'amertume et la teneur en alcool du produit final à partir des autres éléments de recette qui vous sont fournis. Celle-ci n'inclue pas les ingrédients.

Vous êtes libre du choix du langage de programmation, tant qu'il s'agit du langage de codage de développement programmation Python. car c'est le seul nom que le RH connaît. BREW ayant payé un abonnement github, vous y hébergerez tout votre code et vos livrables.

2.2 Description du projet

2.2.1 Données

Les données vous sont fournies . Elles proviennent de Kaggle . Vous devrez documenter ces données. Leur provenance, la description des champs, leurs types, la quantité de données, etc.

2.2.2 Pré-traitement

Les données ne sont pas parfaites, il vous faudra probablement les pré-traiter afin d'éliminer les valeurs manquantes (élimination de lignes ou colonnes). Il faudra aussi éliminer les features redondantes. Ou en créer d'autres qui ont l'air pertinentes.

2.2.3 Analyse exploratoire

Vous tracerez différents graphes de vos données afin d'en obtenir une intuition. Vous pourrez également commencer à entraîner des modèles simples afin d'évaluer la pertinence de ces données.

Tout ce processus devra être documenté. Y compris ce qui ne fonctionne pas. Les maîtres brasseurs sont curieux de tout ce que vous pourrez sortir de ces données.

2.2.4 Choix, entraînement et optimisation d'un modèle

Vous devrez choisir un type de modèle à entraîner sur vos données afin de prédire l'amertume et le taux d'alcool du produit final. Vous êtes experts, votre choix devra être justifié. Vous pourrez également optimiser les hyperparamètres du modèle.

2.2.5 Évaluation des performances du modèle

Un modèle entraîné ne fonctionne pas toujours bien dans la pratique. Vous devrez évaluer ses performances sur de nouvelles données. Si les performances devaient ne pas être bonnes, vous devrez revenir à une étape précédente.

2.2.6 Application finale

L'entreprise BREW n'est pas constituée d'informaticiens. Vous être d'ailleurs le seul. Vous devrez donc fournir une interface web facile d'utilisation pour prédire l'amertume et teneur en alcool d'une recette donnée. Cette interface n'a pas besoin d'être belle, elle a besoin d'être fonctionnelle.

2.3 Planning et livrables

Ce projet est à réaliser sur une durée d'une semaine. Un livrable intermédiaire sera demandé pour vendredi matin.

Il est demandé pour vendredi matin de fournir une documentation décrivant les données, les pré-traitements effectués ainsi que les résultats de l'analyse exploratoire. Vous devrez également fournir le code réalisant ces pré-traitements. Vous fournirez cela via github. Vous mettrez un tag `livrable-intermediaire` sur le commit correspondant au livrable intermédiaire. Vous inviterez le maître brasseur **germain72** ainsi que le RH **SylvainGault** (qui semble avoir pris un intérêt pour votre travail) à vos repositories github.

La livraison finale aura lieu mardi après-midi. Elle comprendra une documentation complétée avec les étapes suivantes. Elle comprendra également le code réalisant la prédiction finale ainsi que le code de l'application web. Le projet devra être présenté lors de la livraison.

2.4 Présentation

Vous présenterez en 30 minutes votre projet. Votre présentation reprendra l'essentiel des données de votre rapport. Celle-ci sera suivie de quelques questions.

Votre présentation devra inclure :

- Le nom des membres du groupe
- Une en-tête avec le logo Diginamic
- Un pied avec date et numéro de page