

# Présentation\_notebook

June 17, 2024

## 1 ml-in-the-clouds

### 1.1 Data Exploration

### 1.2 Préparation de la donnée

- On utilise `TfidfVectorizer` pour transformer des textes en vecteurs numériques pondérés par TF-IDF
- Et `MultinomialNB` pour classer des données textuelles.

### 1.3 Pycaret

#### 1.3.1 Grid Search

### 1.4 Results

```
[1]: ### Results

# 

# 

# 

# ### Results

# 

# 

# 
```

## 1.5 DataRobot

## 1.6 Conclusion

- Performances proches entre Pycaret et DataRobot
- Performance encourageante sur le fichier Train et décevante sur le fichier Test
- La fusion des deux datasets n'a pas été concluante
- Application, Docker fonctionnels
- model déployé sur AWS depuis DataRobot