

# Wikipedia Learning Approach



## Context

UCI Machine Learning Repository:

wiki4HE Data Set

- ☐ Kids
- ☐ Teachers
- ☐ Interesting

## **Problem:**

Is Wikipedia effective in  
students' learning and/or  
teacher's instruction?

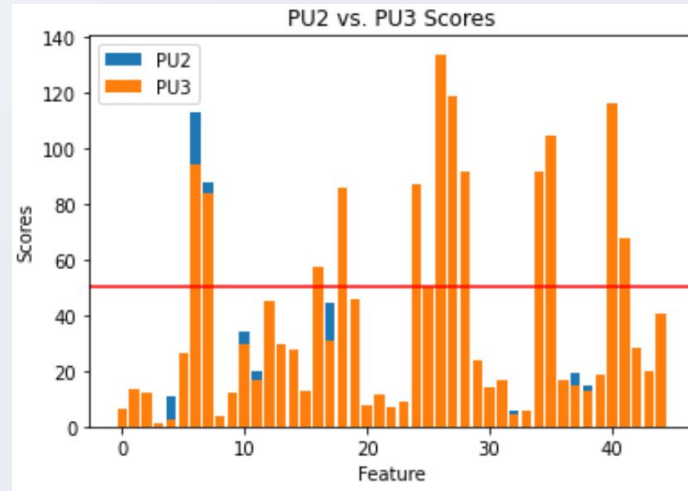
## Dataset

- ❑ Perceived Usefulness
- ❑ Quality
- ❑ Visibility
- ❑ Sharing attitude
- ❑ Use behaviour
- ❑ Behavioral intention

## Response Variable

PU2: The use of Wikipedia improves students' learning

PU3: Wikipedia is useful for teaching

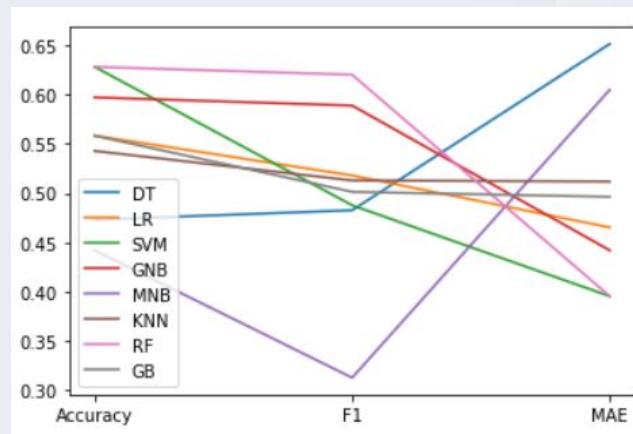
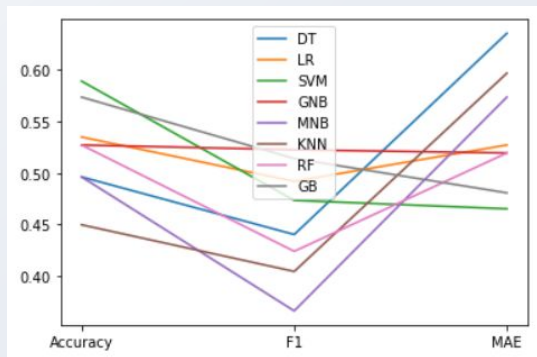


## Models

- ☐ Decision Tree
- ☐ Logistic Regression
- ☐ Support Vector Machine
- ☐ Naive Bayes
- ☐ K Nearest Neighbor
- ☐ Random Forest
- ☐ Gradient Boosting

## Outcomes

- Accuracy
- F1 Score
- Mean Absolute Error (MAE)



# AWS Usage



S3



Data Storage



SageMaker



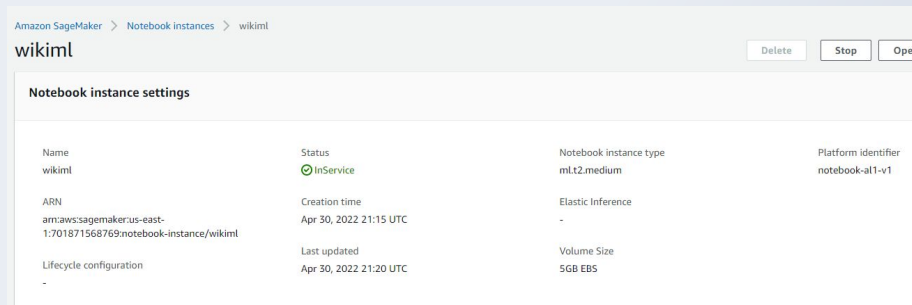
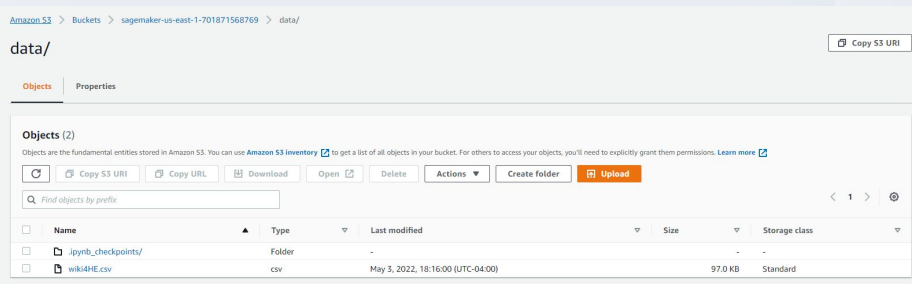
Notebook



Training



Endpoint





## Challenges

- ❑ Choosing Dataset that I am truly interested in
- ❑ Dataset is small
- ❑ Dataset is based only in Spain
- ❑ Choosing the right parameters for each model
- ❑ Making full use of AWS
- ❑ Training the Model

## Future Scope

- ❑ Expand Dataset
  - ❑ Countries
  - ❑ Multiple Universities
- ❑ Adjusting Parameters
- ❑ AWS

# THANKS!

**Any questions?**