

أكاديمية سدایا
SDAIA Academy

**Project report :NLP-DS-Comment and Data science
chat bot**

By

Mohammed alzabyedi

Mukhtar al bin hamad

Bushra Alolyit

Data Science Bootcamp SDAIA Academy

Abstract

The importance of this project comes from the importance of the Data Science as it's the main topics and trendy in these years we choose to make a quick access to any information related to DS fields and make it easy to our users

Design

To help institutions recognize and identify their customers in which field to improve their services.

Algorithms

1. Load dataframe
2. Preprocessing data (removing missing values , unnecassery data,white spaces,numeric values,punctations,stop words,lemmtazition,stemm ,part of speech)
3. Vectorization
4. Topic modeling for 5 component
- 5.Label Comment as AL,ML,DL,DS,and data analysis
6. Exploratory Data Analysis
- 7.Prepare data for modeling
- 8.Classification

Tools

- Numpy and Pandas for data manipulation
- Nltk
- Matplotlib and Seaborn for plotting
- CountVectorizer
- TF/IDF
- Sklearn
- Flask

Data Source

We have used Amazon data science book reviews from [Kaggle](#)

Communication

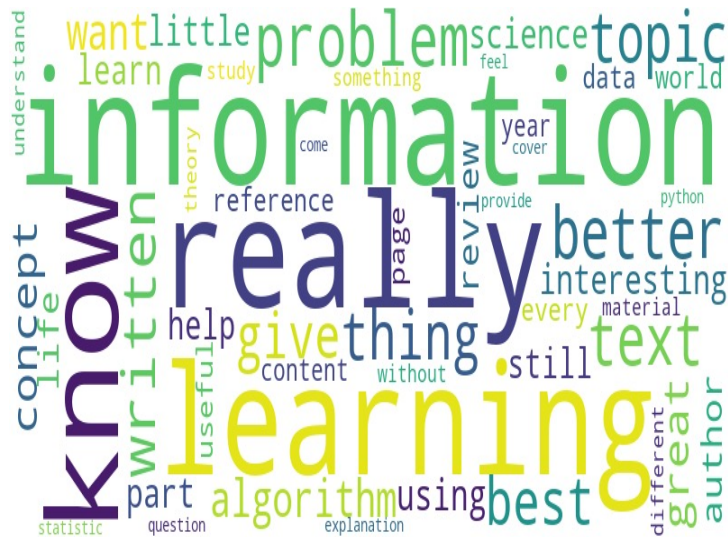
In addition to the slides presented,these are word cloud

○ Data columns

Column	Type
comment	text

The data contain 20647 rows.

○ Word Cloud after cleaning



○ Presentage of Field

