



جامعة عفت  
EFFAT UNIVERSITY

# Aspect-Based Sentiment Analysis of MOOC Learner Reviews

## Machine Learning Final Project

Presented by: Araa, Albatoon, Fahad, and Howyna





# Project Overview

We analyze Coursera learner reviews to extract key aspects and determine sentiment for each.



## Extract Aspects

Identify content, instructor, and platform mentions.



## Analyze Sentiment

Determine positive or negative feelings about each aspect.



## Visualize Results

Present insights through an interactive Streamlit dashboard.

# Problem Statement & Goals

## The Challenge

MOOC platforms receive numerous unstructured reviews.  
Traditional analysis fails to link sentiment with specific  
aspects.

Course designers need refined insights to improve quality.

## Our Goals

- Preprocess raw learner reviews
- Extract aspects using unsupervised techniques
- Classify sentiment for each aspect
- Visualize results in an interactive dashboard
- Evaluate model performance



# Dataset Overview

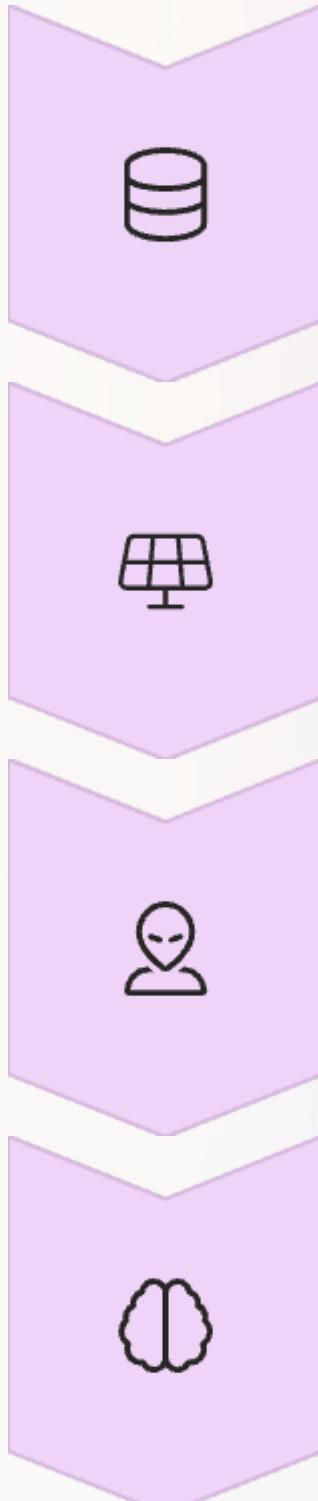
**1.45M**

## Course Reviews

Collected from Coursera

Variable	Type	Description
reviews	Character	Textual content of the review
reviewers	Character	Name of the reviewer
date_reviews	Date	Submission date
rating	Integer	The numeric rating assigned by the reviewer to the course
course_id	Character	The unique identifier associated with the course being reviewed

# Methodology



## Dataset Collection

Coursera MOOC reviews (1.45M reviews)

## Data Preprocessing

Cleaning, tokenization, stopword removal, lemmatization

## Aspect Extraction

Dependency Parsing using SpaCy, RAKE to extract key phrases, and Word2Vec , and Term Frequency & Filtering

## Sentiment

Traditional ML and BERT-based classification

# Train Word2Vec Model for Similarity Filtering

## Aspect Cluster: **content**

- material (0.79)
- information (0.61)
- syllabus (0.59)
- lecture (0.55)
- topic (0.53)

## Aspect Cluster: **video**

- lecture (0.73)
- slide (0.7)
- clip (0.63)
- transcript (0.62)
- podcast (0.61)

## Aspect Cluster: **quiz**

- test (0.76)
- quizz (0.74)
- exam (0.73)
- homework (0.69)
- assessment (0.68)

## Aspect Cluster: **assignment**

- exercise (0.85)
- assigment (0.83)
- homework (0.82)
- assessment (0.76)
- assignement (0.75)

## Aspect Cluster: **teacher**

- instructor (0.83)
- professor (0.82)
- lecturer (0.8)
- tutor (0.77)
- prof (0.66)

## Aspect Cluster: **platform**

- site (0.58)
- opportunity (0.58)
- app (0.55)
- service (0.53)
- environment (0.53)

## Aspect Cluster: **course**

- class (0.65)
- one (0.54)
- program (0.46)
- believe (0.44)
- necessary (0.43)

## Aspect Cluster: **usability**

- modernday (0.82)
- uso (0.82)
- microcourse (0.82)
- weekin (0.81)
- dictatorship (0.81)

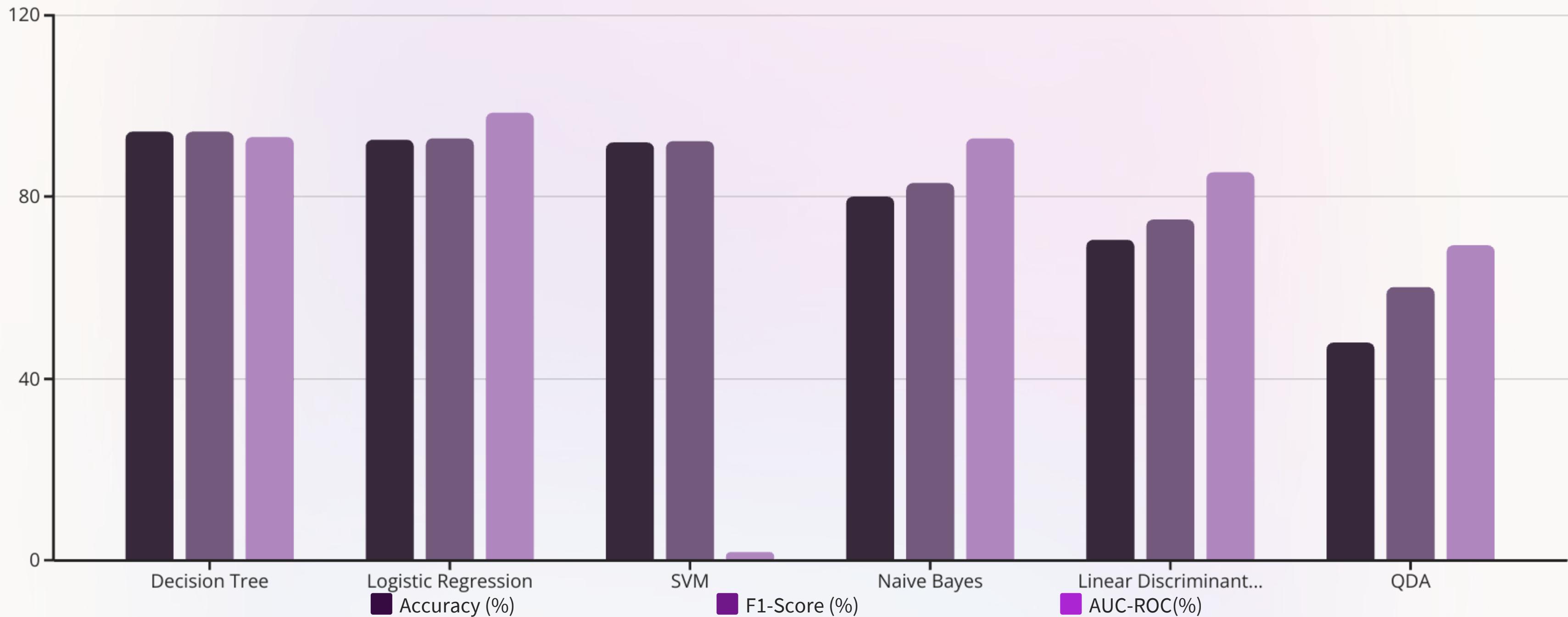
# Aspect Term Frequency



Top 20 Frequent Aspect Terms:

course: 56931  
nan: 30697  
great course: 19505  
good course: 17644  
excellent course: 10834  
good: 10541  
lot: 9087  
thank: 7388  
people: 4711  
amazing course: 4244  
excellent: 4171  
concept: 3772  
nice course: 3597  
great: 3329  
knowledge: 2911  
love course: 2874  
love: 2822  
awesome course: 2626  
thing: 2544  
wonderful course: 2348

# Model Performance Results



Decision Tree outperformed all models with 94.28% accuracy. Logistic Regression had the highest AUC-ROC at 98.45%.

# Dashboard

Fork :

Select Aspect

quiz

assignment

content

course

platform

quiz

teacher

video

## MOOC Aspect-Based Sentiment Dashboard

Analyze learner feedback across aspects like videos, quizzes, instructors, and content.

**Sentiment Distribution for Aspect: quiz**

Sentiment	Percentage
positive	72.1%
neutral	21.4%
negative	6.51%

**Most Discussed Aspects in Reviews**

Aspect	Count
course	~250k
assignment	~20k
video	~15k
content	~10k
teacher	~10k
quiz	~10k
platform	~10k

**Sample Comments for 'quiz' by Sentiment**

Positive Neutral Negative

- ['information perfect program little annoying wait minute watch video take quiz information perfect pass test issue', 'information perfect program']
- ['loss value accord', 'footpark', 'legitimate course concept principle course instructor clear stock market trader quiz answer question stop loss value accord find footpark find stay history time']
- ['disorganized lecture', 'nice professor course many flaw complete focused situation suitable worldwide usersthen disorganized lecture follow logical thread put confused order quiz question unrelated material previous lecture think provide well course old update current event', 'logical thread']
- ['lot content', 'original lecture case', 'lot content seem cut original lecture case question certain quiz relate lesson week future quiz question design compare course coursera', 'certain quiz']
- ['great teacher', 'good solid overview financial market good level depth question quiz appear incongruous internet scavenging aid comprehension peer review midcourse interesting broad good course lecture great teacher']



# Key Learnings

## Unstructured Data

Gained experience processing and extracting meaning from raw text reviews.

## NLP Techniques

Applied advanced natural language processing to connect sentiment with specific aspects.

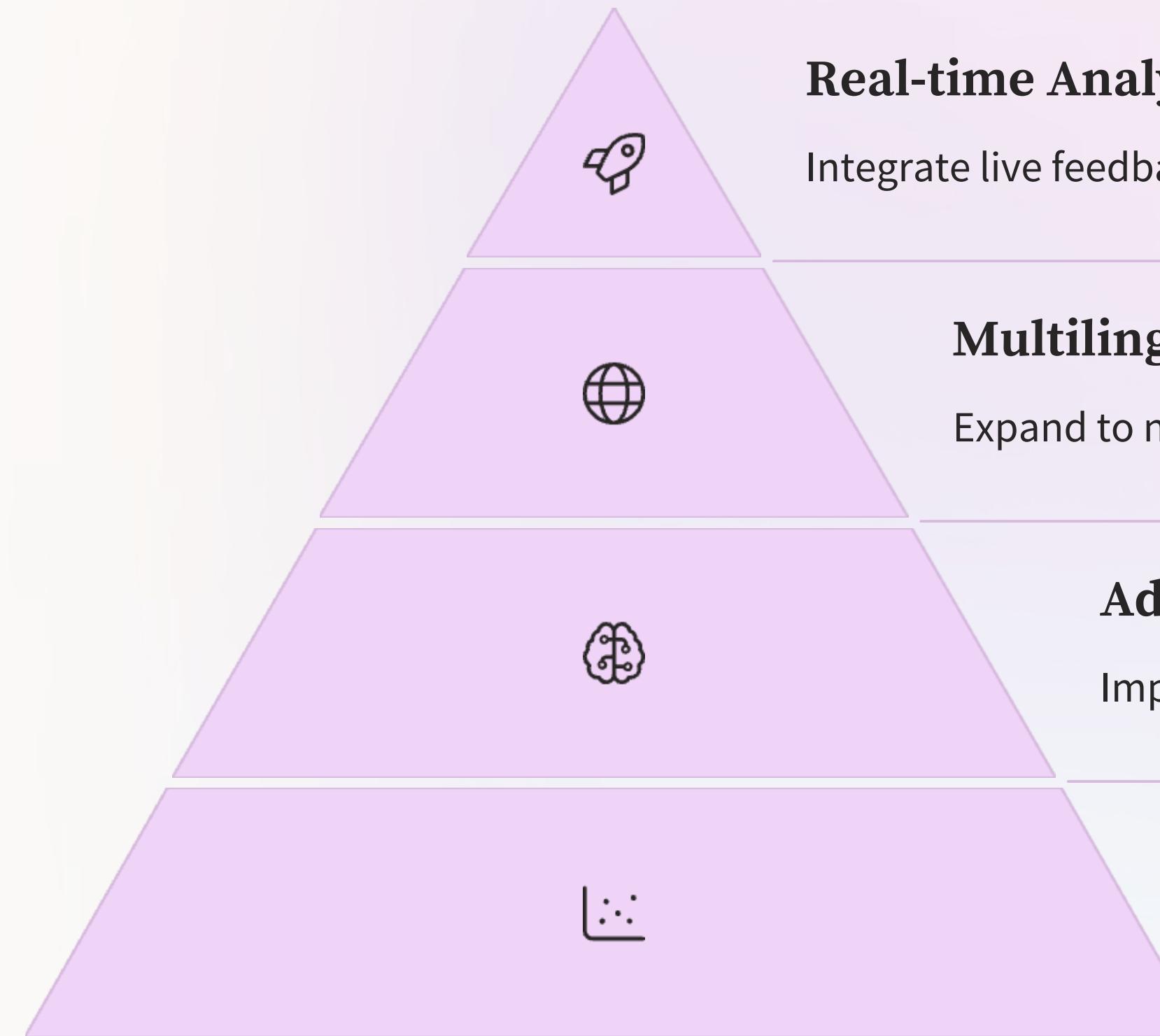
## Model Comparison

Deepened understanding of traditional ML vs. transformer-based models like BERT.

## Performance

Learned to effectively evaluate models using multiple metrics.

# Future Work



## Real-time Analysis

Integrate live feedback processing

## Multilingual Support

Expand to non-English reviews

## Advanced Topic

Implement contextual embeddings

## Enhanced Dashboard

Add interactive visualizations

# Conclusion

## Successful

Built effective aspect-based sentiment analysis system

## Room for Growth

Identified clear paths for future enhancement



## Valuable Insights

Extracted actionable feedback from unstructured reviews

## Practical Tool

Created dashboard for course designers



# Thank You

Any questions?