Mini-Project

Due time (11 May 2024)

Arabic Sentiment Analysis of Hotel Reviews

- **Objective**:
    - The goal of this project is to build and evaluate 3 models for sentiment analysis of movie reviews using Naive Bayes, Logistic Regression, and neural network. You may use python libraries: Numpy, pandas, sklearn, keras, NLTK
- **Data**:
    - The dataset you will use for this project is a collection of text reviews and their corresponding sentiment labels (positive or negative). The review are in Arabic Language.
- **Method**:
    - **Data preprocessing:**
        - Import the dataset and perform basic data cleaning and preprocessing. This includes removing stop words, punctuation, and stemming the text.
        - Split the dataset into training and test sets.
    - **Feature extraction**:
        - Convert the text data into numerical features using techniques such as **bag of words** or **TF-IDF**.
    - **Model building**:
        - Train a Naive Bayes classifier on the extracted features using a training algorithm provided by NLTK library.
        - Train a logistic regression model on the extracted features using a suitable training algorithm.
        - Train a simple neural network model with one hidden layer on the extracted features using a suitable training algorithm.
        -

- **Model evaluation:**
    - Evaluate the performance of the trained Naive Bayes model on the test dataset using metrics such as accuracy, precision, recall, and F1 score.
    - Evaluate the performance of the trained logistic regression model on the test dataset using metrics such as accuracy, precision, recall, and F1 score.
    - Evaluate the performance of the trained neural network model on the test dataset using metrics such as accuracy, precision, recall, and F1 score.
- **Model Comparison:**
    - Compare the performance of the three models and discuss which one performed better and why.
- **Deliverables**:
    - A report detailing your approach and findings, including code and relevant visualizations.

Model	Feat Ext. Method	Acc	Prec	Recall	F1
NB	BoW				
	TFIDF				
LR					
NN					

NN

Epochs	Acc	Prec	Reca	F1	
10					
15					
20					
30					