Arabization Documentation

Project Name: Restaurant Sentiment Analysis

Dr/ Marwa Fekry

Team members:

1- Omar Hussein Mohamed

2- Ali Mohamed Ali

3- Sara Mohamed Ali

4- Nada Hazem Ismail

5- Sama Ayman Ahmed

6- Ahmed Mahmoud Abdelwahab

7- Mazen Mohamed Mahmoud

8- Kareem Abdelfattah Hamed

➤ Definition:

In this project, Arabization means making a sentiment analysis tool that can understand Arabic restaurant reviews well. We want to create a strong app that can figure out if a review is positive or negative. To do this, we need to do a few things like preparing the data, training the computer model, testing it, and then putting it out for people to use. We have to be careful because Arabic has its own unique challenges, like different dialects and cultural differences. So, our goal is not just to understand Arabic text but also to respect the culture and feelings of Arabic-speaking people when they talk about restaurants.

➤ Challenges of Arabic with this APP:

- Preprocessing Complexity: Arabic text often contains diacritics (Taksheel) and special characters, requiring thorough preprocessing to remove noise and standardize text.

- Bidirectional Text Handling: Arabic script is written from right to left, which introduces complexities in text rendering and encoding, Proper handling of bidirectional text is essential to ensure readability and usability of the sentiment analysis application.

- Class Imbalance: Addressing class imbalance in the dataset, where positive and negative reviews may not be evenly distributed, addressing this imbalance is critical to prevent bias in model predictions and ensure accurate sentiment analysis across all review categories.

- Linguistic Variability: Arabic has diverse dialects, regional variations, and linguistic complexities, making it challenging to develop a sentiment analysis model that can accurately understand and interpret sentiment expressions across different Arabic language contexts.

- Cultural Sensitivities: Arabic-speaking communities have diverse cultural norms, sensitivities, and expressions, which may influence the sentiment conveyed in reviews.

➤ Related work:

– " Corpus-based vs. Lexicon-based Approaches": This study compares corpus-based and lexicon-based approaches for sentiment analysis in Arabic restaurant reviews. It explores the effectiveness of each method in capturing the nuances of sentiment expressed in Arabic text.

– "Deep Learning for Arabic Sentiment Analysis in Restaurant Reviews": This research employs deep learning techniques such as recurrent neural networks (RNNs) or transformer-based models like BERT for sentiment analysis of Arabic restaurant reviews. It investigates the performance of these models in capturing sentiment polarity and context in Arabic text.

– "Cross-Lingual Sentiment Analysis for Arabic Restaurant Reviews": This study explores the challenges and techniques involved in adapting sentiment analysis models trained on other languages to analyze Arabic restaurant reviews. It may investigate approaches like translation, cross-lingual transfer learning, or domain adaptation to overcome language barriers and cultural nuances in sentiment analysis.

– "Arabot": Arabot is chatbot platform focuses on users that speak Arabic. Sentiment analysis is used to interpret user messages and provide suitable emotional intelligence responses. It can be included into a number of platforms such as social media, messaging applications, and websites to successfully interact with people, respond to questions, and offer customer care.

➤ Future work:

– Scaling for Real-time Analysis: As the volume of online restaurant reviews continues to grow, scaling the sentiment analysis application to handle real-time analysis of large datasets from diverse sources is essential. Implementing efficient data processing pipelines and distributed computing architectures can enable timely and scalable sentiment analysis of Arabic restaurant reviews.

- <u>Extending to Other Text Domains</u>: Beyond restaurant reviews, there are numerous other domains where sentiment analysis of Arabic text can be valuable, such as product reviews, news articles, social media conversations, and customer feedback surveys. Extending the sentiment analysis application to analyze sentiment in these diverse text domains can broaden its applicability and utility across various industries and use cases.

- <u>Social Media Monitoring Tool for Arabic Restaurants:</u> Building a tool that monitors social media platforms for Arabic restaurant reviews and sentiment analysis can provide valuable insights into customer opinions and trends. By tracking mentions of specific restaurants or keywords related to dining experiences, this tool can help restaurant owners and managers understand customer sentiment over time and make informed decisions to improve their services.

- <u>Arabic Restaurant Review Sentiment Analysis Chatbot:</u> Developing a chatbot that engages with users in Arabic and provides sentiment analysis for restaurant reviews offers a convenient way for users to access sentiment analysis results and restaurant recommendations. Users can interact with the chatbot to inquire about specific restaurants, ask for sentiment analysis of reviews, or explore sentiment trends in the restaurant industry.