Service Quality AI Based System

Supervised by : Dr Ahmed Salah , T.A Moataz Mohamed

ABSTRACT

The concept of smart government has been gaining momentum in recent years, as governments around the world strive to improve service delivery and enhance citizen engagement with technology. In this context, the Smart Government Service Quality Project represents an innovative and ambitious initiative aimed at improving the quality of government services and enhancing citizen satisfaction. Our project has applied the target desired from smart government service quality idea as we have deployed many AI NLP-based techniques as topic classification, sentiment analysis, text generation and information retrieval, these techniques allowed us to analyze reviews, complains and feedbacks of citizens in order to get desired outputs automatically without need of human participating, these outputs includes: the ministry responsible for the review, the sentiment of the citizen posting the review, the specific governmental sector responsible for the problem in the review and the recommended solution for the problem in the review. We have deployed a website to make it easy for every citizen to use it from his house, work, or any place with the internet. We have applied fine tuning for transformer models like MARBERT , Arabert, Arat5, MT5 on our downstream tasks, we have reached **0.96** f1 score in sentiment analysis task, **0.957** f1 score in topic classification task and **7.8** BLEU score in text generation task.

CONTACT

Amr ayman
Ahmed samir

Youssef mohammed
Mostafa gamal
Ali fathy
Zyad el sayed

Ain shams
amr2018170267@cis.asu.edu.eg

INTRODUCTION

The Smart Government Service Quality Project is an innovative initiative aimed at improving the quality of government services through the use of advanced technologies.
The project leverages data analytics, artificial intelligence, and other smart technologies to optimize government service delivery processes, reduce bureaucracy, enhance transparency, and increase citizen satisfaction.
The project is based on a customer-centric approach that puts citizen needs and preferences at the center of service delivery. Through this approach, the project seeks to transform the way governments deliver services, making them more efficient, effective, and responsive to the needs of citizens.

METHODS AND MATERIALS

WE HAVE CREATED THIS PROJECT WITH MANY SCRIPTS, WE USED PYTHON PROGRAMMING LANGUAGE AS IT WAS THE SIMPLEST AND MOST USED LANGUAGE IN THE FIELD OF AI BECAUSE IT IS OPEN SOURCE, EASY LANGUAGE AND HAVE ENORMOUS NUMBER OF PACKAGES AND LIBRARIES THAT ARE VERY USEFUL IN OUR PROJECT FIELD.
ENVIRONMENTS:
LOCALHOST ,GOOGLE COLAB,GOOGLE COLAB PRO ,KAGGLE NOTEBOOKS:SAME AS GOOGLE COLAB WITH ADDITIONALTPU.
PACKAGES & LIBRARIES
PANDAS,.MATPLOTLIB,OS,.NUMPY ,PY TORCH ,TRANSFORMERS ,SENTENCE PIECE,TQDM:,JSON,TENSORFLOW:,SCIKITLEARN,NLTK,RE,JOBLIB,SELENIU, BEAUTIFUL SOUP.
DATASET PREPROCESSING
GLARE DATASET PREPROCESSING,MERGED TOPICS DATA PREPROCESSING, WEB SCRAPPING MODULE, INPUT TEXT PREPROCESSING MODULE.
TOPIC CLASSIFICATION MODULE:
MODELS USED: MARBERT,ARABIC-MARBERT-NEWS-ARTICLE-CLASSIFICATION
DATA PREPARATION: CORPUS COLLECTION,TEXT PREPROCESSING.
KEY PREPROCESSING TECHNIQUES INCLUDE: TEXT CLEANING,TOKENIZATION,LEMMATIZATION AND STEMMINGFEATURE EXTRACTION,LABEL ENCODING.
TRAINING PROCEDURE: SPLITTING THE DATASET,MODEL SELECTION ,MODEL TRAINING,HYPERPARAMETER TUNING.
EVALUATION METRICS: WE USE METRICS SUCH AS MACRO F1 SCORE.
SENTIMENT ANALYSIS MODULE:
MODELS USED: MARBERT MODEL ,ARABERT MODEL.
HERE IS A DESCRIPTION OF THE FUNCTIONS TYPICALLY EMPLOYED IN ARABERTFORSENTIMENTANALYSIS:
DATAPREPARATION,MODEL ARCHITECTURE,LOSSFUNCTION,OPTIMIZATIONALGORITHM,TRAININGLOOP,EVALUATIONMETRICS ,HYPERPARAMETER TUNING.
RECOMMENDED SOLUTION GENERATION MODULE:
MODELS USED:GOOGLE\MT5,UBC\NLP ARAT5,AUBMIND\LAB ARAGPT
DATA PREPARATION:
THE PREPROCESSED INPUT TEXT IS TOKENIZED USING THE MODEL'S TOKENIZER, THEN WE EXTRACT THE FEATURE USING THE TRANSFORMER MODEL REPRESENTATION TO BE ABLE TO BE PASSED TO THE ARCHITECTURE OF THE MODEL.
TRAINING PROCEDURE:FINE TUNING,TRAINING LOOP,LOSS FUNCTION,OPTIMIZATION ALGORITHM
EVALUATION METRICS:WE USED BLEU SCORE FOR EVALUATION AS IT IS A METRIC THAT MEASURES THE SIMILARITY BETWEEN A MACHINE-GENERATED TEXT AND ONE OR MORE HUMAN-GENERATED REFERENCE TEXTS.
INFORMATION RETRIEVAL:
TECHNIQUES USED:TF-IDF REPRESENTATION,SENTENCE COSINE SIMILARITY ALGORITHMS:FEATURE EXTRACTION,SIMILARITY CHECK,RANKING.

RESULTS

Topic classification
After large number of experiments using different models, applying different text preprocessing, and applying hyperparameter tuning we concluded that experiment 7 using marbert model and preprocessed Merged topics dataset gives the best f1 score as shown in table 4.

Sentiment Analysis
After large number of experiments using different models, applying different text preprocessing and applying hyperparameter tuning we concluded that experiment 5
using arabert model and preprocessed combination between 100k reviews dataset and glare dataset, gives the best f1 score as shown in table 5.

Text Generation
After large number of experiments using different models, applying different text preprocessing and applying hyperparameter tuning we concluded that experiment 3
using arat5 model and preprocessed glare dataset, gives the best bleu score as shown in table 6.

Information Retrieval
Here are some outputs from our algorithm shown in figure 38, figure39 .

Table 4.

Experi ent No.	Catego ries	Train Data size	Test Data size	Batch Size	Learn ing Rate	Mode l used	Score	Datas et
1	7	8000	2000	32	1e-06	MAR BERT	f1 : 0.15	Glare
2	7	8000	2000	32	1e-05	MAR BERT	f1 : 0.75	Glare
3	7	16000	2000	32	5e-05	MAR BERT	f1 : 0%	Glare
4	4	8000	2000	32	2e-06	MAR BERT	f1 : 0.01	Glare
5	4	8000	2000	32	2e-05	MAR BERT	f1 : 0.55	Glare
6	4	8000	2000	32	5e-05	MAR BERT	f1 : 0.447	Glare
7		4400	1100	32	5e-05	MAR BERT	f1 : 0.956	Major Topics

Table 5.

Experi ent No.	Catego ries	Train Data size	Test Data size	Batch Size	Learn ing Rate	Mode l used	Score	Datas et
1	3	4000	1000	32	2e-06	MAR BERT	f1 : 0.15	Oclar
2	3	16000	4000	32	5e-05	MAR BERT	f1 : 0.655	Glare
3	3	8000	2000	32	5e-05	MAR BERT	f1 : 0%	100k revs
4	2	8000	10000	32	2e-06	MAR BERT	f1 : 0.71	Glare
5		15600	2000	16	2e-05	Araber	f1 : 0.96	100k glare

Table 6.

Experie ent No.	Catego ries	Train Data size	Test Data size	Batch Size	Learn ing Rate	Mode l used	Score	Datas et
1	-	4000	1000	5	1e-04	Mt5 small	bleu : 0.1154	Glare
2	-	16000	4000	10	5e-04	Arat5	bleu : 7.047	Glare
3	-	80000	20000	13	5e-05	Arat5	bleu : 7.8017	Glare
4	-	60000	10000	8	1e-03	Gpt2 arabic	bleu :	Glare
5	-	150000	15000	8	1e-03	Gpt2 arabic	bleu :	Glare

Figure 38.

المشكلة : تلوث الهواء من كثره الدخانات والثلوثات الكثير
سنزل هذا المكره ال القطاعات العامه ب : البيئة
الحل التي يجب ان يتم : تعزيز استخدام الطاقة المتجددة والتكنولوجيات النظيفة

Figure 39.

المشكلة : نقص الادويه في المستشفيات
سنزل هذا المكره ال القطاعات العامه ب : الصحة
الحل التي يجب ان يتم : تحسين جودة الخدمات الصحية في المستشفيات الحكومية
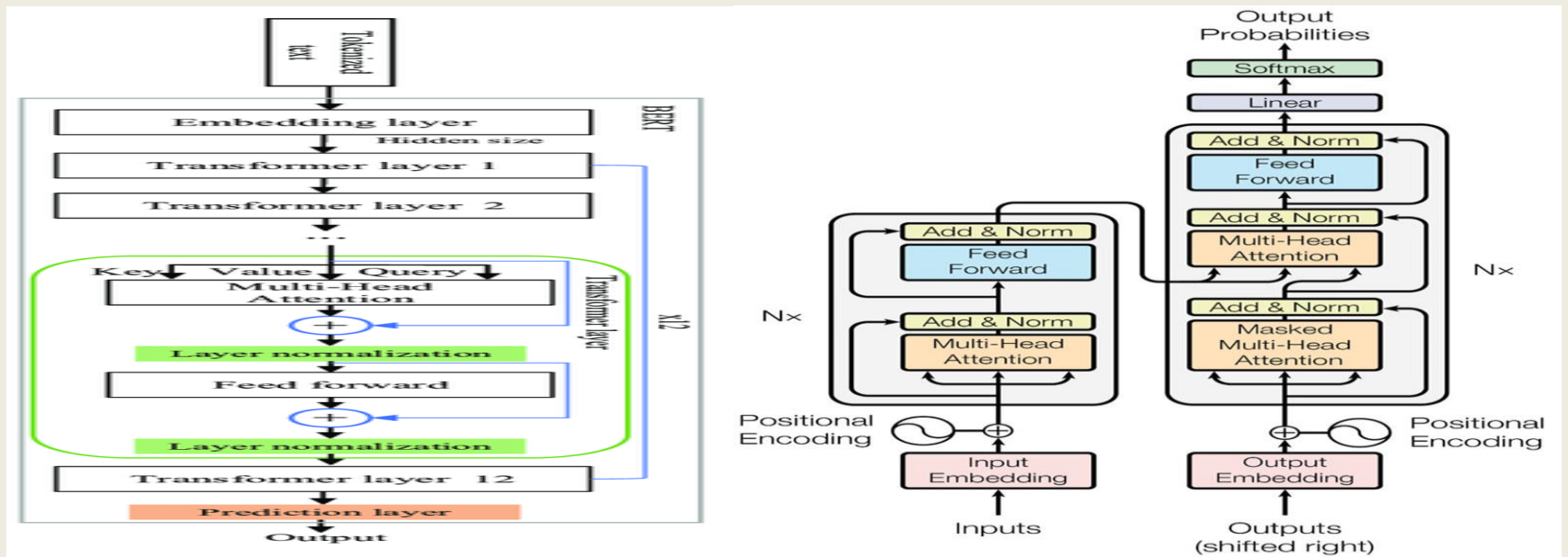
Acknowledgements

CONCLUSIONS

The Smart Government Service Quality Project is an important initiative aimed at improving the quality of government services through the use of smart technologies. By leveraging data analytics, artificial intelligence, and other advanced technologies, this project can help governments to identify and address service delivery bottlenecks, reduce bureaucracy, enhance transparency, and increase citizen satisfaction. With its focus on continuous improvement and innovation, this project has the potential to transform the way governments deliver services, making them more efficient, effective, and responsive to the needs of citizens.The project can be able to automatically analyze the text reviews of citizens and get the topic(ministry) of the review, the sentiment of citizen writing the review, the governmental sector responsible for the review content and recommend solutions for the problems in the review.In this project we reached best scores in topic classification with **0.957** f1 score, in sentiment analysis with **0.96** f1 score, in generating the solution and sector for the review with reasonable solutions coming from the database of known services for the government.

REFERENCES

[1]Abdul-Mageed, M., Elmadany, A. and Nagoudi, E.M.B., 2020. ARBERT & MARBERT: deep bidirectional transformers for Arabic. arXiv preprint arXiv:2101.01785. [2]Antoun, W., Baly, F. and Hajj, H., 2020. Arabert: Transformer-based model for arabic language understanding. arXiv preprint arXiv:2003.00104.