FACULTY OF ENGINEERING

REAL TIME FEEDBACK: CASE OF CROP DISEASES DIAGNOSIS APPLICATION FOR SMALLHOLDER FARMERS

By

Omara Jonathan

Reg.No BU/UG/2017

Bachelors of Computer Engineering

Department of Computer and Electrical Engineering

omarajonathan@gmail.com , +256785762653

SUPERVISOR

Dr. Owomugisha Godliver

Department of Computer and Electrical Engineering

ogodliver@gmail.com, +256774820997

Co -SUPERVISOR

Ms. Asingwire Barbara Kabwiga

Department of Computer and Electrical Engineering

kezabarbara@gmail.com, +256774618076

A project implementation report submitted to the department of Electrical and Computer Engineering in partial fulfillment to the requirements for the award of a Bachelor's Degree in Computer Engineering at Busitema University.

APPROVAL

Student details

Name: **Omara Jonathan**

Signature ……

Date ……

Supervisor

Name: **Dr. Owomugisha Godliver**

Signature ……

Date ……

Co-supervisor

Name: **Ms. Asingwire Barbara Kabwiga**

Signature ……

Date ……

DECLARATION

I ..... earnestly declare that I am the sole writer of this report and it has never existed before in any Institution of higher learning or University for the same award. The information in this report is my personal effort and not a duplicate of any proposal report.

Signature .....

Date .....

ABSTRACT

Uganda is one of the countries in Africa which has an agro-based economy that feeds the population and other foreign countries through exports. Proper crop disease diagnosing tools and information about agricultural practices is key to the growth and development in the agricultural sector. In order to help farmers, get real time answers on queries about agronomic practices and also get accurate diagnosis about their crops, we have built an android application using a dataset collected from farmers in eastern Uganda and some parts of northern Uganda. This application can answer queries on land preparation, seed selection, planting, crop maintenance, harvesting and good storage practices on any of the three major crops in Uganda that's cassava, maize and beans. It can also help farmers diagnose cassava crop diseases with an accuracy of 94%. This application is active and available 24/7. The Q&A component was built using machine learning and natural language processing. In building the model, an S. BERT transformer model was used and even with a Bleu score accuracy of 60%, the system provides useful and clear feedback to agronomic queries in any of the three crops mentioned above. With such an application, farmers can progress towards easier information about farming related practices, crop disease diagnosis and hence a better agricultural output.

Table of Contents

APPROVAL I

DECLARATION II

ABSTRACT III

1 CHAPTER 1: Introduction 1

1.1 Background 1

1.1.1 Agriculture in Uganda 1

1.1.2 A Smallholder farmer 1

1.1.3 Machine learning 2

1.2 Problem Statement 2

1.3 Objectives 3

1.3.1 General objectives 3

1.3.2 Specific Objectives 3

1.4 Significance of the Study 3

1.5 Scope 3

2 CHAPTER 2: Literature Review 4

2.1 Existing Agricultural applications in Uganda 4

2.2 Existing Applications in other Countries and NLP techniques 5

2.2.1 Existing in other Countries 5

2.2.2 NLP Techniques applied in building real time feedback systems 5

3 CHAPTER 3: Methodology 8

3.1 Data collection 8

3.2 Feature Extraction 8

3.2.1 Data analysis 8

3.2.2 Stop words removal 8

3.2.3 Stemming 9

3.2.4 Lemmatizing 9

3.2.5 TFID Vectorization 9

3.3 Building and Training Real Time Feedback Models 9

3.3.1 Transfer Learning 9
3.3.2 Retrieval based-model 9
3.3.3 Generative Models 12
3.3.4 Ensemble based model 13
3.4 Testing of the model accuracy 13
3.4.1 Precision and recall (F1 score) 13
3.4.2 **Bilingual Evaluation Understudy(**BLUE) 14
3.4.3 ROUGE 14
3.4.4 **Perplexity** 14
3.4.5 Integration into a mobile App 14
4 CHAPTER 4: Methods and Materials 15
4.1 Data Collection 15
4.1.1 Interviewing farmers 15
4.1.2 Online database 16
4.2 Text Preprocessing 18
4.2.1 Spelling Correction and punctuation removal 18
4.2.2 Rephrasing answers 18
4.2.3 Removing unanswered questions 18
5 CHAPETR 5: Implementation and Metrics 19
5.1 Mobile Application Components 19
5.1.1 Image classification task (Diagnostic tool) 19
5.1.2 QA task (Feedback tool) 19
5.2 NLP Model Selection 21
5.2.1 S. BERT model Architecture 23
5.3 Results 24
5.4 Metrics 25
5.5 Model Deployment and app integration 26
6 CHAPTER 6: Discussion and Conclusion 27
6.1 NLP Model and Android app 27

6.2 Recommendations 27
6.3 Conclusion 27
Appendix A: Attached CD containing code for the entire project 28
Appendix C: References 29

List of Acronyms

NLP	Natural language processing
Seq2seq	Sequence to Sequence
TFID	Term Frequency Inverse Document
GRU	Gated Recurrent Unit
LSTM	Long Short Term Memory
RNN	Recurrent Neural Network

List of figures

Figure 1: Diagnostic tool UI 7
Figure 2: Block diagram for the proposed system..... 7
Figure 3: Seq2Seq Architecture [37] 13
Figure 4: Map of Uganda showing the green spots of data collection 16
Figure 5: Districts sending the most amount of data 17
Figure 6: Chart showing the distribution of questions in each of the three crops 17
Figure 7 UI: Image classification task.....20
Figure 8 UI: QA task.....20
Figure 9: Table showing model performance comparison.....21
Figure 10: Graph showing comparison between execution time for each model tested on same 10 questions 22
Figure 11: S. BERT model architecture.....23
Figure 12: Blue score for 10 questions with and without punctuations 26