West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
La Paz, Iloilo City

# EXCEIPT: AUTOMATED BOOKKEEPING ENCODING THROUGH IMPROVED INFORMATION EXTRACTION ON RECEIPTS

An Undergraduate Thesis

Presented to the Faculty of the

College of Information and Communications Technology

West Visayas State University

La Paz, Iloilo City

In Partial Fulfillment

of the Requirements for the Degree

Bachelor of Science in Computer Science

by

Dave F. Fagarita

Jimuel S. Servandil

Jannica Mae G. Magno

Jeziah Lois C. Catanus

June 2024

West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
La Paz, Iloilo City

Approval Sheet

EXCEIPT: AUTOMATED BOOKKEEPING ENCODING THROUGH IMPROVED INFORMATION EXTRACTION ON RECEIPTS

An Undergraduate Thesis for the Degree

Bachelor of Science in Computer Science

by

Dave F. Fagarita

Jimuel S. Servandil

Jannica Mae G. Magno

Jeziah Lois C. Catanus

Dr. Arnel N. Secondes
Panel

Dr. Ma. Luche P. Sabayle
Panel

Dr. Frank I. Elijorde
Panel

Nerilou B. Dela Cruz
Panel

John Cristopher A. Mateo
Adviser

Concurred:

Dr. Ma. Luche P. Sabayle
Chair, Computer Science

Dr. Ma. Beth S. Concepcion
Dean

June 2024

# Acknowledgment

The researchers would like to extend their heartfelt gratitude to the following people who helped them accomplish everything in conducting their study:

First and foremost, to God Almighty, for unconditional love and for His unending mercy. He had showered countless blessings upon the researchers and had bestowed them wisdom and will to continue the journey;

Mr. John Cristopher Mateo, thesis adviser and Ms. Nerilou Dela Cruz, thesis co-advisor, for contributing their knowledge, ideas, inspiration, and motivation to complete this study and the guidance and support provided from the beginning to the finish of their struggles. They give their time to review the paper and constructively critiquing it for improvement are very much wanted;

The participants, for providing their time despite their hectic schedules in order to share their experiences, which considerably contributed to the study's findings. Their participation will be a great inspiration to those who planned to apply for a similar job.

The researchers' families, friends, and support systems provide unending moral, emotional, and financial support at every step, comforting the researchers during times of failure.

The Iloilo Grace Pharmacy, WVSU Cooperative, and 7-Eleven San Joaquin and Leon Branch, for willingly providing us the receipts that were vital for training our model.

The Supply Office and Mr. Topaz from Accounting for their valuable contributions, cooperation, guidance and assistance.

Dave Fagarita

Jimuel Servandil

Jannica Mae G. Magno

Jeziah Lois C. Catanus

June 2024

Fagarita, Dave, Servandil, Jimuel, Magno, Jannica Mae, Catanus, Jeziah Lois; "ExCeipt: Automated Bookkeeping Encoding through Improved Information Extraction on Receipts". Unpublished Undergraduate Thesis, Bachelor of Science in Computer Science, West Visayas State University, Iloilo City, Philippines, June, 2024.

## Abstract

Efficient bookkeeping is essential for business operations, but manually encoding receipts can be time-consuming and error-prone. This study develops an automated system to extract structured data from receipts using machine learning, streamlining the bookkeeping process. The proposed system utilizes OCR and LayoutLMv3 model, a state-of-the-art pre-trained model for document layout analysis and information extraction. A Convolutional Neural Network (CNN) algorithm validates receipt authenticity prior to processing. The methodology involved data collection of 498 receipts, annotation, preprocessing, model fine-tuning, evaluation, and implementation.

The LayoutLMv3 model achieved 99.12% accuracy and 95.35% F1 score in extracting and labeling receipt information. The CNN model for receipt validation attained 99.77% accuracy. The system was integrated into a web

application, enabling users to upload receipts and obtain structured merchant details, transaction dates, items, prices, and taxes. Evaluated using ISO/IEC 25010 standards, the system demonstrated high quality and effectiveness across the 8 categories. This automated system enhances bookkeeping efficiency and accuracy, benefiting bookkeepers, small-to-medium enterprises, and stakeholders in accounting and finance domains.

Table of Contents

# List of Figures

# List of Tables

# List of Appendices

# CHAPTER 1 INTRODUCTION TO THE STUDY

## Background of the Study and Theoretical Framework

Keeping purchase invoices or receipts is crucial for businesses, freelancers, and households when it comes to bookkeeping. These records serve as evidence of purchase and contain vital information like the date, amount, and type of transaction. Maintaining these records is essential for expense tracking, cash flow monitoring, and financial statement preparation (Hayes, 2022; Ionos, 2018).

However, managing receipts and inputting their contents into the processing software is a laborious and time-consuming job for bookkeepers. Failing to handle accounting and bookkeeping accurately can lead to serious compliance issues for any company (Vakilsearch, 2023). To avoid any delays in bookkeeping tasks, it is crucial to streamline the flow of transactions and receipt generation.

Machine learning algorithms and rule-based systems are used in Information Extraction to extract important data from unstructured text. This information is stored in a database, automating receipt encoding and speeding up bookkeeping (Ontotext, 2022).

Boston Consulting Group suggests that automating processes is necessary to meet the growing demand for faster and more efficient bookkeeping. Implementing this technology poses challenges (Brackert et al., 2019). This study aims to develop an automated system that can extract structured information from receipts. Automating this task allows bookkeepers to focus on other important financial processes, saving time in the long run (Libeo, 2022). The researchers present this study in order to determine the most efficient way to automatically extract receipt, as well as to assess the level of accuracy and effectiveness of the automation process. The most effective method for automating receipt encoding can be identified, enabling bookkeepers to concentrate on important financial tasks instead of encoding receipts (Tipalti, 2023).

Theoretical Framework

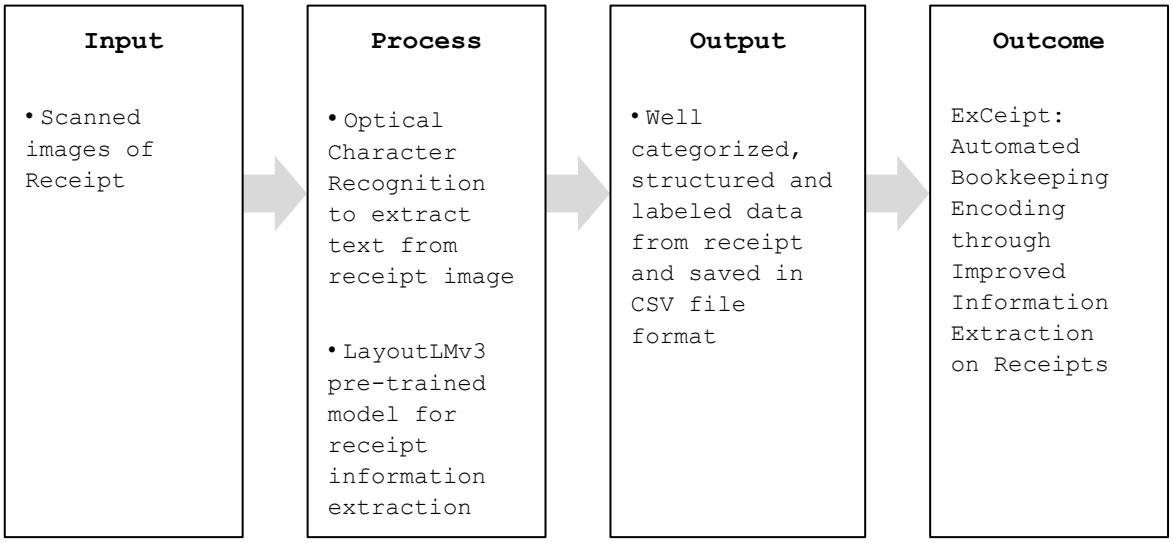



**Figure 1.** Theoretical Framework

Figure 1 provides an overview of the automated process to extract structured data from receipts and encode it for bookkeeping. The diagram consists of stages for input, process, output, and outcome. As input, receipts in English-language are included, along with detailed lists of items purchased, their quantities, prices, and dates of the transactions. For receipt information extraction, detection, and layout analysis, the LayoutLM pre-trained model is used, along with a machine learning model to enhance the data extraction accuracy. The output includes accurate structured data from the receipt, such as the

receipt number, merchant name, merchant address, date and time, items, prices, total amount, and vat tax, which can be automatically encoded into bookkeeping data and entered into the accounting system. The outcome of the process is improved receipt information extraction in the automated system, which leads to more accurate and efficient financial reporting and analysis.

## Objectives of the Study

The study sought to achieve its general objective of creating an automated bookkeeping encoding system through improved information extraction on receipt.

To achieve this goal, the study had identified several specific objectives such as;

1) To implement a Convolutional Neural Network algorithm to validate if the image is a receipt or not.

2) To fine-tune the LayoutLMv3 model to the receipt dataset for document image understanding and layout analysis and utilize the model to automatically recognize the category of the text and assign it to its respective category.

3) To evaluate the performance of the model trained on a receipt dataset and use the F1 metric for text labeling tasks.

4) To develop a web application that scans and extracts the information from the receipt.

5) To create a structured tabular data extracted from the receipts and evaluate its performance using the standards of ISO-IEC 25010 System Product Quality Evaluation Tool.

## Significance of the Study

The study has numerous benefits for different stakeholders, which include:

*Bookkeepers*

Automating receipt encoding can help bookkeepers save time and boost productivity. This will enable them to concentrate on other crucial duties and optimize their time usage.

*Small to Medium sized Businesses*

Reduced labor costs and fewer errors can benefit these businesses, avoiding costly rectification. Automating receipt encoding can improve data accuracy and compliance for better decision-making.

*West Visayas State University*

The success of this study could enhance the university's reputation as a leading research institution. Contribution to the field of accounting and finance is possible.

*Government Agencies*

Automated receipt encoding can help government agencies monitor business transactions and ensure tax law

compliance. It can also assist in reducing the possibility of fraud and tax evasion.

*Future researchers*

The findings of this study can be a valuable reference or foundation for future research. This can enable individuals to expand their existing knowledge and contribute to the creation of innovative ideas and solutions.

## Definition of Terms

The section below defines key terms used in this study conceptually and operationally.

*Automated* - refers to tasks that are performed by machines or computers without requiring human intervention (Cambridge, n.d).

In this study, the system "Automated" automatically extracts and structures information from receipts.

*BERT* - is a pre-trained language model developed by Google researchers (Lutkevich, 2020).

*Bookkeepers* - accounting professionals who are responsible for maintaining detailed records of financial transactions (MyAccountingCourse, 2023).

In this study, the system will be used by bookkeepers to manage accounting and bookkeeping records for a business.

*Convolutional Neural Networks* - an Artificial Neural Network used to recognize and process images (DeepAI, 2019).

In this study, Convolutional Neural Networks are used to process structured data arrays, like images.

*CSS (Cascading Style Sheets)* - is a programming language that is utilized to manage the appearance and design of web pages (Mozilla Foundation, 2023).

In this study, HTML, CSS and JavaScript are the web frameworks used in the development of the web app.

*Encoding* - is the process of converting data into a format suitable for various information processing requirements (Techopedia, 2022).

In this study, encoding refers to inputting extracted receipt information into a structured database.

*Flask* - Flask is a micro web framework for Python that facilitates the development of web applications by providing a simple and lightweight structure (Makai, 2022).

In this study, Flask will be used as the backend framework of our web application.

*HTML (Hypertext Markup Language)* - is a markup language that is utilized to produce web pages (Shneiderman, 2019).

*Hugging Face Transformer* - is a library of pre-trained natural language processing models (Wolf et al., 2020).

In this study, Hugging Face Transformer is used as the base foundation for training the layoutLM in our specified dataset.

*Information Extraction* - is the procedure of retrieving particular data from written sources (Ontotext, 2022).

In this study, Information Extraction refers to the process of extracting information from receipts and saving it in a structured format.

*JavaScript* - is a programming language that is commonly used to add interactivity to web pages (Mozilla Foundation, 2023).

*LamBERT* - is a pre-trained language model developed by Facebook AI. (Devlin et al., 2020)

*LayoutLMv3* - The model for deep learning that was utilized to carry out document processing (Nanonets, 2022).

In this study, LayoutLMv3 will be utilized for analyzing and comprehending the layout of receipts.

*Optical Character Recognition* - processing tool for transforming digital images of physical documents into a format that can be read by machines (TechTarget, 2022).

In this study, Optical Character Recognition is used to convert text into code for data processing.

*Python* - A programming language used for website and software development, task automation, and data analysis (Coursera, n.d). In this study, the web application will be developed using Python as the main programming language.

*Receipt* - is a written record of a payment made for goods or services received (Sweta, 2023).

In this study, receipts were used as input for the web application and data for training and testing the model.

*Recordkeeping* - is the process of creating and maintaining records of business activities (American Institute of Certified Public Accountants, 2023).

In this study, only the RecordKeeping is the primary focus of the study and the blueprint of the features used in the web application.

*RoBERTa* - is a pre-trained language model developed by Facebook researchers (Liu et al., 2019).

In this study, the three pre-trained models namely BERT, RobERTa, and LambERT are the most popular pretrained models that are currently used in hugging face websites.

*Small to Medium firms (SMEs)* - are businesses with fewer than 250 employees (Small Business Administration, 2023).

*UBIAI* - A user-friendly text annotation tool specifically designed for NLP tasks, making it ideal for efficient labeling (UBIAI, 2023).

In this study, it will be used as an annotation tool for manually labeling the receipts.

## Delimitation of the Study

This section enumerates the delimitations that are present in the study. This includes the following:

- The study will only include scanned taped transactions and acknowledgment receipts that are written in English.
- The study will only focus on the primary step or RecordKeeping of the bookkeeping process.
- The study will only use LayoutLMv3 in pre-training the model for document image understanding and layout analysis.
- The data used in the study will be anonymized and will conform to data privacy laws.
- The study will focus only on small to medium-sized firms or establishments.
- The study will gather data from receipts owned by researchers, as well as receipts voluntarily provided by others and private businesses that are willing to share their receipts.
- The study will only extract the necessary information on the receipt, such as the Merchant Name, Merchant Address, Transaction Date, Transaction Time, Items, Price, Total and Vat Tax.

The delimitation of this study serves to define the boundaries and scope of the research focused on automating receipt extraction and encoding in order to improve bookkeeping. It ensures a clear path for study and research goals within a specific framework.

# CHAPTER 2 REVIEW OF RELATED STUDIES

## Review of Existing and Related Studies

This chapter provides a comprehensive literature review on receipt processing and bookkeeping automation. It provides an overview of current research and technical achievements in the area, with a particular emphasis on strategies for extracting and encoding essential information from receipts to improve the efficiency and precision of bookkeeping duties.

Automated receipt processing and bookkeeping have become increasingly popular in recent years. There are various methods available to automate bookkeeping, such as utilizing accounting applications to monitor income, expenses, invoice statuses, and quarterly as well as annual taxes (Myers, 2021). Another method involves connecting payment processor accounts to an accounting application to automatically monitor income (Myers, 2021). Moreover, the utilization of artificial intelligence (AI) has enabled the automation of financial document processing, including invoices and receipts (Cabuz, 2019).

*Studies related to Bookkeeping and Accounting*

According to Vickneswaran (2016), accounting is described as the core activity of any organization. It involves the collection, analysis, and dissemination of financial data pertaining to economic entities. Bookkeeping, on the other hand, refers to the act of recording financial transactions and is an integral part of the accounting process in every organization. Transactions typically encompass a range of activities such as purchases, sales, receipts, and payments conducted by either individuals or organizations (Vickneswaran, 2016).

As mentioned by the study of Vickneswaran (2016), organizations have the option to efficiently handle their accounting and bookkeeping tasks using Microsoft Office's Excel sheet. This solution is not only cost-effective but also user-friendly. However, it is crucial for the organization to customize the accounting and bookkeeping system to suit their particular business type and meet their financial statement requirements. While accounting software offers greater efficiency compared to manual accounting and bookkeeping, it is essential to possess sufficient knowledge for its effective operation.

Furthermore, acquiring accounting software can also pose a substantial financial commitment for the company.

*Traditional Bookkeeping and Accounting Methods*

To successfully manage income and spending and record financial transactions, people have relied on traditional bookkeeping and accounting techniques for ages. Using spreadsheets or paper ledgers, individuals are typically required to manually record transactions using these old methods. This process is frequently error-prone and time-consuming. These methods can be useful, but they frequently take a long time and are prone to mistakes.

To improve the accuracy and effectiveness of bookkeeping, new tools and methods have been developed recently-thanks to technological breakthroughs. Accounting software is one such technology that may automate several processes, including creating reports, calculating taxes, and keeping track of income and expenses. According to IBM (2022), using OCR technology can save time, cost, and other resources by utilizing automated data extraction and storage capabilities.

Another tool that has been used to improve bookkeeping is optical character recognition (OCR) technology. OCR can be used to scan paper receipts and automatically extract relevant information such as the date, vendor name, and amount. This information can then be imported into an accounting program for easier tracking. ZenBusiness Inc. (2022) notes that OCR technology has greatly impacted how businesses handle documents with accounting being one area that has benefited from this.

Essentially, while modern technologies like accounting software and OCR technology have their benefits in terms of accuracy and efficiency, old bookkeeping techniques still have their place but it lacks convenience, efficiency and reliability.

*Machine Learning and Deep Learning*

The processing of receipts and bookkeeping have both been automated using machine learning and deep learning techniques. On this subject, several researchers have been done. One such study is by UiPath, which has developed an AI game-changer called UiPath Receipt and Invoice AI. Developed by UiPath’s world-class machine learning (ML) and

AI team, this new AI activity enables UiPath Robots to read both invoices and receipts and helps automate accounts payable and expense compliance processes (Cabuz, 2019).

Another study by Sudharshan Chandra Babu discusses using OCR (Optical Character Recognition) and Deep Learning for automating receipt digitization. The article covers the theory behind receipt digitization and implements an end-to-end pipeline using OpenCV and Tesseract (Babu, 2022).

Microsoft has also developed a pre-built model called Receipt processing prebuilt AI model. This pre-built model uses state-of-the-art optical character recognition (OCR) to detect printed and handwritten text and extract key information from receipts (Fernandez et al., 2023).

*Related Studies on Automated Receipt processing*

Huang, Chen, et al. (2019) provide a comprehensive report on the challenge, including the competition datasets, task definitions, evaluation protocols, submission statistics, and an analysis of the submitted performance. The challenge results revealed that text localization and recognition tasks were fairly straightforward to handle, whereas the task of information

extraction posed more difficulties. It indicates that there is still opportunity to improve the efficiency of the information extraction task despite the various ideas and methods that have been presented forth.

Xu, Zhang, et al. (2020) underline how crucial it is to recognize how information extraction and text reading interact when interpreting document images. They emphasize the significance of understanding the connections between these two processes. Their suggested method demonstrates the efficiency of a well-coordinated approach for this particular task and has the potential to raise both the accuracy and the efficiency of systems made to comprehend document images.

Sun, Kuang, et al. (2021) put forward a comprehensive approach for extracting key information from unstructured document images. Their method is designed to overcome the shortcomings of traditional template matching techniques, which face difficulties when dealing with document images that have unfamiliar templates and are prone to errors in text recognition. However, it is crucial to conduct additional research to evaluate the effectiveness of this

method on more diverse datasets and in scenarios that extend beyond the scope of the current study.

Bardelli, Rondinelli, et al. (2020) demonstrate how electronic invoicing may be used to create a smart system that automates the accounting process by categorizing invoices into distinct codes. Machine learning algorithms are effective in this task, and considering the hierarchical structure of the account codes can improve performance. This research has implications for the efficiency and accuracy of accounting processes for businesses in Italy and beyond.

*LayoutLM*

In the papers of Xu, Li et al. (2020), a pre-training model named LayoutLM is proposed to address or solve the limitations of the current pre-training models in NLP that only focus on text-level manipulation and ignore important layout and style information. By incorporating layout information, LayoutLM greatly improves tasks related to document image understanding, such as form and receipt comprehension, as well as document image classification, resulting in state-of-the-art outcomes. Furthermore, the

model integrates visual data by incorporating image features during its pre-training phase. It is worth mentioning that this particular model is the pioneer in simultaneously acquiring knowledge on both text and layout within a cohesive framework for document-level pre-training.

In their study, Bunch (2021) showcases the practical implementation of the LayoutLM model in extracting data from scanned receipts. The LayoutLM model is a deep learning model that has been trained extensively on a large corpus of text and layout information. This training equips the model to excel in tasks like Optical Character Recognition (OCR) and extracting information from scanned documents.

The article provides a detailed tutorial on how to adjust the LayoutLM model using the SROIE dataset, which is designed for OCR and extracting information from scanned receipts. The authors present a step-by-step guide on preparing the data, configuring the model, and training and evaluating it on the SROIE dataset.

According to the authors, the LayoutLM model, after being fine-tuned, has shown remarkable accuracy in extracting important details like the total amount, date, and vendor name from scanned receipts. In addition, they conducted a comparison between their model and other advanced models, demonstrating that their approach surpasses existing methods.

The article focuses on showcasing the impressive abilities of deep learning models such as LayoutLM in effectively extracting valuable information from scanned documents. It serves as a valuable resource for researchers who are keen to delve into this particular field of study.

*Related Studies on Applications of LayoutLM*

Ku, Lei et al. (2020) introduced the LayoutLM- critic, a new iteration of the LayoutLM concept. With a primary focus on aligning text and images, this model is specifically created to handle visually rich materials such as scanned papers and PDFs. The model goes through a pre-training task called the "masked visual-text alignment" task to do this. By matching the bounding box of the text segment with the appropriate image throughout this task,

the model gains the ability to anticipate the location of a text segment within an image. The model must correctly pinpoint the location of the masked text inside the image after being exposed to photos during the training phase that contain masked text segments.

According to Verma L. (2022), the LayoutLM model presents an innovative method that integrates text and layout information for handling scanned document images. This combined approach has showcased its advantages in various practical tasks related to document image analysis, particularly in the extraction of information from scanned documents.

In addition, LayoutLM utilizes image features to effectively incorporate visual information from words into its framework. This groundbreaking approach combines text and layout information within a single framework for document-level pre-training, which to the best of our knowledge is a first. The integration of these elements has led to remarkable improvements in performance across a range of downstream tasks, such as form understanding, receipt comprehension, and document image classification.

As a result, LayoutLM has established new state-of- the-art results in these domains.

*LayoutLM Version 2*

Xu et al. (2022) presented LayoutLMv2, an innovative approach to pre-training models for tasks involving the comprehension of documents that contain visual elements. This method integrates text, layout, and image information within a unified framework. It incorporates a spatial-aware self-attention mechanism into the Transformer architecture to effectively capture the relationships between various bounding boxes. Moreover, new pre-training objectives are employed to facilitate the model's understanding of how different modalities interact with each other.

*LayoutLM Version 3*

The paper presents LayoutLMv3, a study conducted by Huang, Lv et al. (2020), a novel pre-training model for multimodal Transformers in Document AI. The model aims to enhance existing approaches by using a combined text and image masking objective, in which a word-patch alignment objective is employed to predict if the corresponding image patch of a text word is masked, thus enabling cross-modal

alignment. The proposed approach simplifies the architecture and training objectives and results in a versatile pre-trained model suitable for both text-centric and image-centric Document AI tasks. The study demonstrates that LayoutLMv3 outperforms existing models and achieves state-of-the-art performance in several tasks, including form understanding, receipt understanding, document visual question answering, document image classification, and document layout analysis. The code and models used in this work are publicly accessible.

*LAMBERT, BERT, and RoBERTa Model*

*The Bert Model*

Xu, Li et al. (2020) define BERT or Bidirectional Encoder Representations from Transformers as an attention-based bidirectional language modeling approach. The BERT model has been shown to effectively transfer knowledge from self-supervised tasks with large-scale training data. Its architecture consists of a multi-layer bidirectional Transformer encoder that takes in a sequence of tokens and uses multiple layers to produce final representations. The tokens are processed using WordPiece, and input embeddings

are generated by adding the word, position, and segment embeddings. The input embeddings then go through a multi-layer bidirectional Transformer that has an adaptive attention mechanism to create contextualized representations.

*LAMBERT Model*

The LAMBERT model is a layout-aware language model that creates contextualized token embeddings for tokens encountered in formatted documents, according to Garncarek et al. (2021). According to the study, LAMBERT performs better than the industry standard RoBERTa at extracting information from visually complicated documents while retaining performance on papers with a more straightforward appearance. The ability to train in an unsupervised environment, a major strength of language models, is preserved in the model, enabling it to make use of the enormous volume of publicly accessible unannotated documents.

*RoBERTa Model*

Liu et al. (2019) conducted a comprehensive study exploring various design choices in the pretraining of BERT models. Their findings indicate that there are several factors that can significantly enhance performance. These factors include extending the duration of pretraining, using larger batch sizes and more training data, eliminating the next sentence prediction objective, training on longer sequences, and dynamically modifying the masking pattern during training.

To tackle these design challenges, the authors propose a novel pre-training approach known as RoBERTa. The results obtained by the authors underscore the significance of these often overlooked design decisions and demonstrate that BERT's pre-training objective remains competitive when compared to other recently proposed alternatives. This research contributes to the growing body of literature on pre-training methods for language models and sheds light on the potential benefits of incorporating these design choices in future studies.

*Gaps in Previous Studies*

*BERT Model*

Bidirectional Encoder Representations, or BERT for simple terms, is a natural language processing approach that has found major acceptance in a variety of applications. Although it has many benefits, there are a number of disadvantages to using it as well.

The immense size of BERT is one of its main drawbacks. Due to its huge training dataset and training approach, the model is quite massive. Since there are so many weights that need to be changed, the training procedure might be laborious. Additionally, BERT is expensive to use due to its size, which can be a major burden for those integrating it into their business.

BERT's principal function as an input for other systems, rather than as a standalone application, is another limitation. As a result, it demands careful fine-tuning for downstream operations. There is no assurance that the best outcomes will be obtained after this fine-tuning step, which frequently demands more time and resources.

*LAMBERT Model*

According to a study conducted by Garncarek et al. (2021), there is a possibility that enhancing the performance of LAMBERT for visually rich documents can be achieved by increasing the size of the unsupervised training dataset. The authors also express their intention to investigate the impact of enlarging the model size and training datasets to achieve even greater improvements.

Expanding the size of the model has the potential to improve performance as it allows the model to capture more complex patterns and relationships in the data. However, this approach presents practical difficulties such as requiring more computational resources and longer training times.

Expanding the training dataset can also enhance the model's ability to learn from a diverse set of examples, which may result in better generalization performance. However, acquiring and annotating large datasets can be a time-consuming and costly process. Furthermore, it is essential to carefully address ethical concerns related to data privacy and bias.

The authors' intention to explore larger model size and more extensive training datasets is in line with the typical approach in machine learning research to advance performance in a specific task. Nevertheless, it is essential to find a balance between these endeavors and take into account practical considerations and ethical implications.

*RoBERTa Model*

RoBERTa, which stands for Robustly Optimized BERT Pre Training Approach is an incredibly powerful model used in natural language processing. It has the capability to excel in various applications. Nevertheless, it is crucial to acknowledge that RoBERTa has certain limitations and potential drawbacks that need to be taken into account.

One major drawback of RoBERTa is its high computational cost, which can pose challenges for smaller organizations or those with limited computing resources. The model's vast number of parameters makes training and utilization difficult. Additionally, the significant energy consumption of the model raises environmental concerns.

Another drawback of RoBERTa is that it has not been trained on domain-specific language or jargon. As a result,

it may have lower accuracy when dealing with specialized fields like medicine or law. Furthermore, the model's comprehension of broader context might be limited, which can impact its performance in tasks such as question-answering or summarization.

*Summary*

Document image understanding is a task of great complexity in natural language processing, particularly when it involves documents that possess intricate layouts and visual components. To address this challenge, the LayoutLM model has been developed, which integrates layout and style information with text manipulation techniques.

The researchers plan to use LayoutLMv3, the third version of the LayoutLM model and the same pre-trained model that has been applied in other studies, in the suggested investigation. This model performs better in tasks like form understanding, receipt understanding, and document picture classification because it integrates image features with text and layout learning. The model can handle documents with complex layouts, but there are still certain limits, necessitating more development.

To overcome these limitations, more recent iterations of LayoutLM have been developed, including LayoutLMv2 and LayoutLMv3. These versions integrate additional pre-training objectives and spatial-aware self-attention mechanisms, enhancing the model's capability to comprehend documents with visual elements. These advancements significantly improve the accuracy and robustness of LayoutLMv2 and LayoutLMv3 when it comes to tasks related to understanding document images.

In addition to LayoutLM, other pre-training models like BERT, LAMBERT, and RoBERTa have gained widespread adoption. While BERT is effective in transferring knowledge, it can be computationally expensive and slow. On the other hand, LAMBERT outperforms RoBERTa in extracting information from visually complex documents, but it would benefit from larger training datasets. RoBERTa, despite its drawbacks such as computational expense and limited understanding of domain-specific language and broader context, is still widely used.

Liu et al. (2019) conducted a study that proposes specific design choices to improve the effectiveness of

pre-training methods in language models. These decisions include task-specific pre-training, utilizing more training data, and fine-tuning on specific downstream tasks. Incorporating these design decisions into future work could significantly improve the performance of pre-training models like LayoutLM and its variants.

While the LayoutLM model and its newer versions have demonstrated improved performance in document image understanding tasks by incorporating layout and style information, there are still gaps in its capabilities, particularly in handling documents with complex layouts. Addressing these challenges and adapting the model to new domains or tasks will be the focus of the proposed system.

Compared to other pre-training models such as BERT, LAMBERT, and RoBERTa, LayoutLM has its own strengths and limitations. Future research can build upon these models by incorporating the suggested design decisions to enhance their performance in document image understanding and other NLP tasks. With the most recent version of the LayoutLM, the LayoutLMv3 has addressed many issues such as accuracy and the efficiency of the model. In conclusion, the ongoing

development and fine-tuning of pre-training models like LayoutLMv3 will contribute to the advancement of natural language processing and enhance the researchers document image understanding.

# CHAPTER 3 RESEARCH DESIGN AND METHODOLOGY

## Description of the Proposed Study

The main goal of the proposed study is to create and test an automated system that can extract and encode information from receipts, as well as improve the bookkeeping processes. To achieve this objective, the study employs various methods, such as data collection, pre-processing, algorithm development, evaluation, optimization,and lastly, implementation.

In the methodology process, the researchers must first collect a dataset of taped transaction and acknowledgment receipts with standardized formats and layouts. The dataset is crucial as it serves as the training data for the algorithm.

The receipts will then be pre-processed to prepare them ready for information extraction. It involves converting receipt images into machine-readable text using OCR technology and accurately streamlining the program.

Algorithm development follows the preprocessing step. The analysis incorporates deep learning techniques, such as convolutional neural networks (CNN) for receipt

classification, and uses LayoutLMv3 to extract text in the receipt. Moreover, using data enhancement techniques such as noise or distortion incorporating images is used to improve and increase the diversity of the training data.

The algorithms that were developed will then be evaluated to assess their accuracy and efficiency in extracting the information from receipts. The evaluation process will involve comparing the extracted information with the existing models in order to determine the accuracy of the algorithm.

Once the evaluation process is complete, the algorithms are further optimized to enhance their performance. The optimization process may include fine-tuning the algorithm by using information from the analysis or adding further improvements.

And lastly, the optimized algorithms will be implemented into a bookkeeping system to automate the encoding of receipts. The implemented system will be thoroughly tested to ensure that it is accurate and efficient in practice.

## Methods and Proposed Enhancements

### *Methodology*

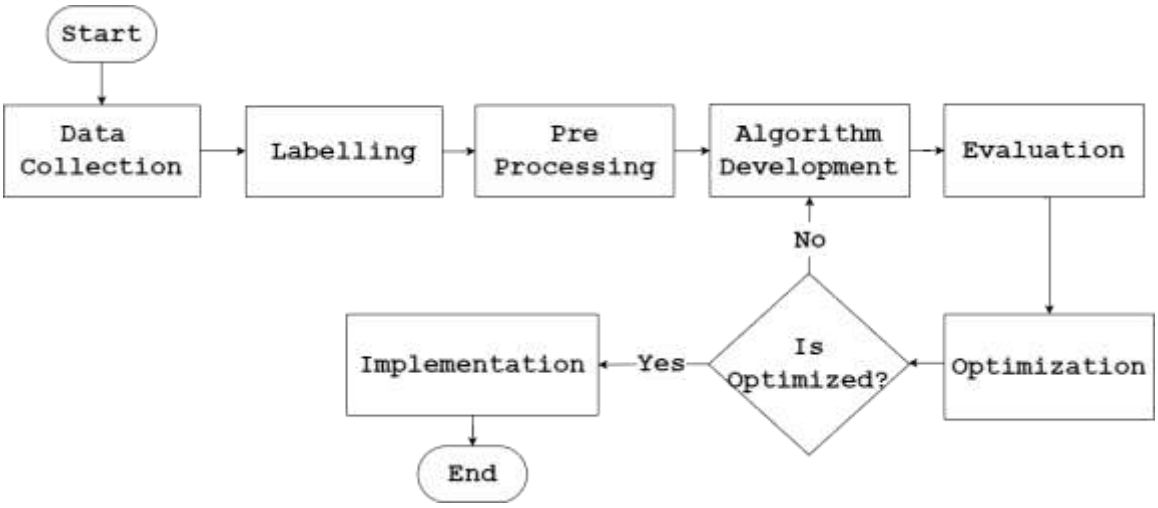



**Figure 2.** Methods Structure

The figure illustrates the typical machine learning workflow, consisting of the following steps: Data Collection, Data Labeling, Data Preprocessing, Algorithm Development, Model Evaluation, and Optimization. If the evaluation metrics are unsatisfactory, the process loops back to Algorithm Development for further refinement. Once the model performance is acceptable, the final algorithm is implemented.

*Data Collection*

Dataset collected was gathered from different establishments like 7 Eleven, Grace Pharmacy, and WVSU Cooperatives. There are almost 1200+ total receipts collected from the said establishments and the data collected are scanned for further processing. However, it comes down to 498 receipts since there are constraints in the annotation tool that the researchers have used.

*Labeling*

The collected taped transaction and acknowledgment receipts are then labeled, which involves manually annotating the selected area of information to be extracted from the receipts and correcting the text extracted by the OCR. The researchers utilized UBIAI. It returns a OCR processed file containing the bounding box, the OCR extracted texts, entities values. This step is crucial for the next step of the process.

*Pre-processing*

The preprocessing of receipts for information extraction involves several important steps, with file

conversion being one of the initial processes. Converting the OCR processed from UBIAI studio into a file format that LayoutLMv3 recognizes is an essential step in preparing the receipts for further analysis.

The preprocessed data undergoes steps that involve data cleaning and data normalization. Data cleaning will perform actions such as identifying errors, inconsistencies, and missing values in the receipts. This will make sure that the data are accurate and reliable for further analysis. After the researchers perform data cleaning, they perform data normalization to standardize the collected data. By converting the data into a standard scale or range, this technique makes it possible to compare and analyze the complete dataset in a meaningful way.

*Algorithm Development*

In the process of information extraction, the algorithm development phase begins after the researchers preprocess the taped transaction and acknowledgment receipts. The LayoutLMv3 model, a customized deep learning model for document layout analysis and information extraction, is being fine-tuned throughout this phase. In

order to extract important data from the preprocessed receipts, such as the merchant's name, merchant address, transaction dates, transaction times, item descriptions, prices, total, and VAT Tax, the pre-trained LayoutLM model must be modified. To improve the accuracy and performance of the model, the model is trained on the preprocessed dataset and optimized accordingly.

Aside from fine-tuning the LayoutLM model, the researchers also implemented convolutional neural networks (CNNs) in the algorithm development phase. This algorithm will validate if the image being supplied to the web application is a valid receipt or not. By utilizing these deep learning techniques, the researchers hope to improve information retrieval accuracy and scope.

*Evaluation*

The accuracy and efficiency of the developed algorithms in extracting information from receipts need to be evaluated. This evaluation would require comparing the extracted information with manually entered data to determine the algorithms' accuracy. The researchers will

use evaluation metrics such as F1 Score, Recall, and Precision.

To calculate the F1 Score, they can use this formula:

$$F1\ Score\ =\ 2\ \frac{Precision\ *\ Recall}{Precision\ +\ Recall}$$

In calculating the Recall, they can use the following formula:

$$Recall\ =\ \frac{No.of\ True\ Positives}{No.of\ True\ Positives\ +\ No.of\ True\ Negatives}$$

To calculate the Precision, they can use the following formula:

$$Precision\ =\ \frac{No.of\ True\ Positives}{No.of\ True\ Positives\ +\ No.of\ False\ Positives}$$

*Optimization*

Optimizing algorithms plays a pivotal role in the development process as it significantly contributes to improving their performance in terms of accuracy and efficiency. Fine-tuning the models, adjusting hyperparameters, or refining the algorithms' logic allows for iterative enhancements that align with the desired

objectives. By fine-tuning the models, it can iteratively improve the algorithm's behavior; ensuring it addresses potential weaknesses or biases present in the initial design. Adjusting hyperparameters helps in finding an optimal configuration that maximizes performance metrics. Moreover, refining the algorithms' logic involves revisiting and improving the underlying structure and implementation, enabling them to address inefficiencies, bottlenecks, or inaccuracies. This iterative optimization process empowers them to enhance the algorithm's overall efficiency and accuracy, making it more robust, reliable, and effective for a wide range of applications.

*Implementation*

The final step would be to implement the optimized algorithms into a flask web application to automate the encoding of receipts. The implemented system would then be tested to ensure its accuracy and efficiency in practice.

## Proposed Enhancements

This study can be further enhanced through the integration of advanced techniques and methodologies. One proposed enhancement is the utilization of LayoutLMv3, a pre-trained language model specifically designed for layout analysis on documents, including receipts (X. Li, 2019). By leveraging LayoutLMv3, the model can effectively capture and understand the intricate layout information present in receipts, leading to improved information extraction accuracy.

In addition to LayoutLMv3, the application of deep learning techniques, such as Convolutional Neural Networks (CNN), can significantly enhance the performance of information extraction from receipts (Y. LeCun, 2013). These models have demonstrated remarkable capabilities in learning complex features and patterns within both the textual and layout components of receipts, contributing to improved accuracy and efficiency.

Furthermore, incorporating domain-specific knowledge can provide valuable context and reduce potential errors during the information extraction process (Ontotext, 2022).

By doing some interviews in the accountant, bookkeeper, and supply office, the researchers can have a better understanding of the anatomy of a News.

Implementing a more advanced Optical Character Recognition (OCR) tool significantly enhances the model's performance by refining the initial stages of training and testing the text extraction process. A sophisticated OCR tool is pivotal in accurately converting images or scanned documents into machine-readable text, thereby improving the quality and reliability of data input for the model. The precision and efficiency of this preliminary step directly impacts the subsequent stages of training and testing, as the model relies on accurate and well-extracted textual information.

These proposed enhancements have the potential to significantly advance the accuracy and efficiency of automated bookkeeping encoding through improved information extraction on taped transactions and acknowledgment receipts.

Components and Design

*System Architecture*

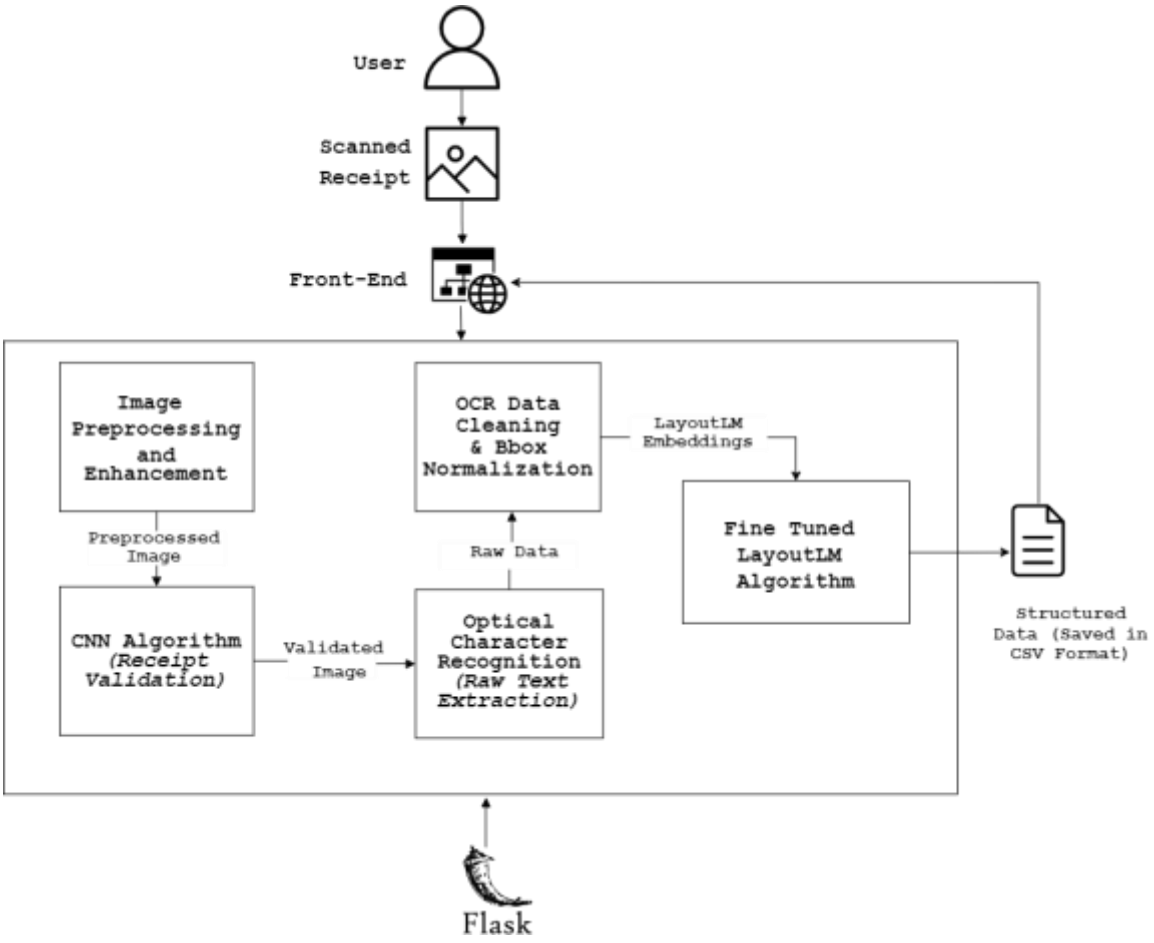



**Figure 3.** ExCeipt Web App System Architecture

As shown in Figure 3, the front-end interface of the system architecture provides a user-friendly platform for users to input their receipts. The interface captures the image of the receipt and sends it to the system for processing. The system uses several image processing

techniques to prepare the image for OCR, including bounding box normalization, data cleaning, and image enhancement, which improves the image quality and reduces noise. The image will then be validated for receipt classification. These models learn what links the different elements of the receipt. If the validated receipt is done, the system then extracts the text from the image by passing the preprocessed image through the OCR. which will be analyzed using the fine-tuned LayoutLMv3 model. The LayoutLMv3 model categorizes the receipt's layout, including any tables, lists, or free-form text, and extracts pertinent information.

After that, the front-end interface receives the structured data to be reviewed by users. The system architecture offers a quick and easy way to process receipts and extract pertinent data, reducing the need for human data entry and improving the data's accuracy.

*Procedural and Object-Oriented Design*

Procedural design describes how a program will be implemented, organized and structured to perform a particular task. It involves breaking down the program into

a set of smaller, discrete procedures or subroutines that are executed sequentially to accomplish a specific goal.

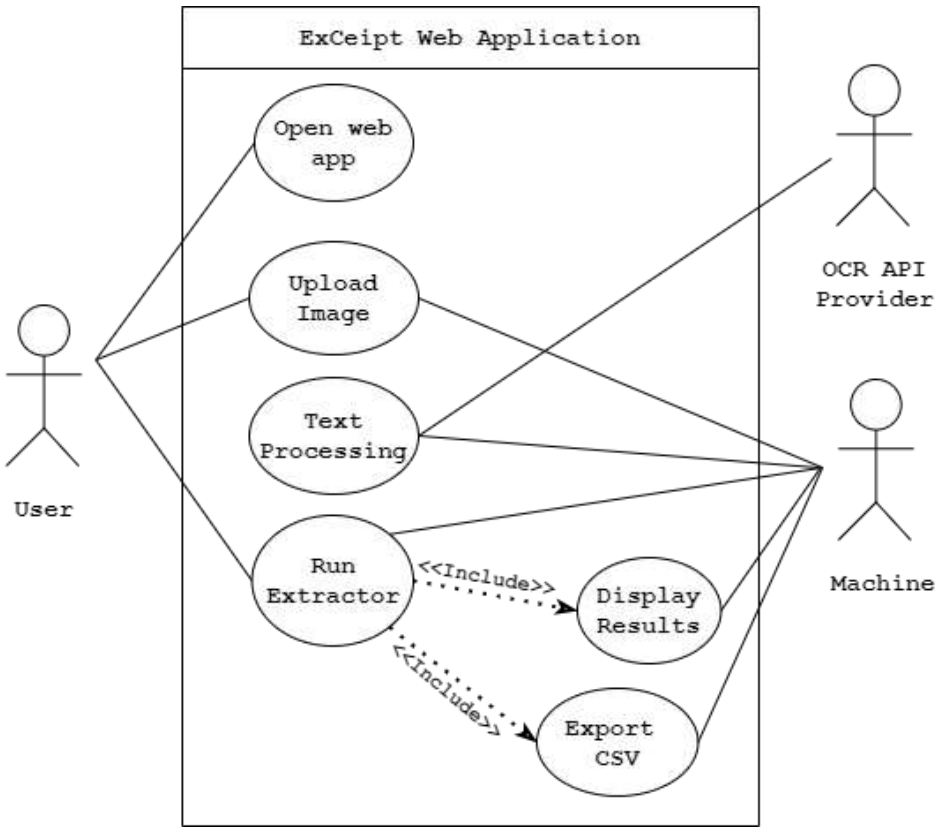



**Figure 4.** ExCeipt Web App Use Case Diagram

In Figure 4, it illustrates the various use cases of our system. Users interact by opening the app, uploading images, and utilizing the extractor function for insights. Meanwhile, the OCR API Provider manages text processing alongside a dedicated machine, which handles image uploads, executes the extractor, displays results, and exports data.

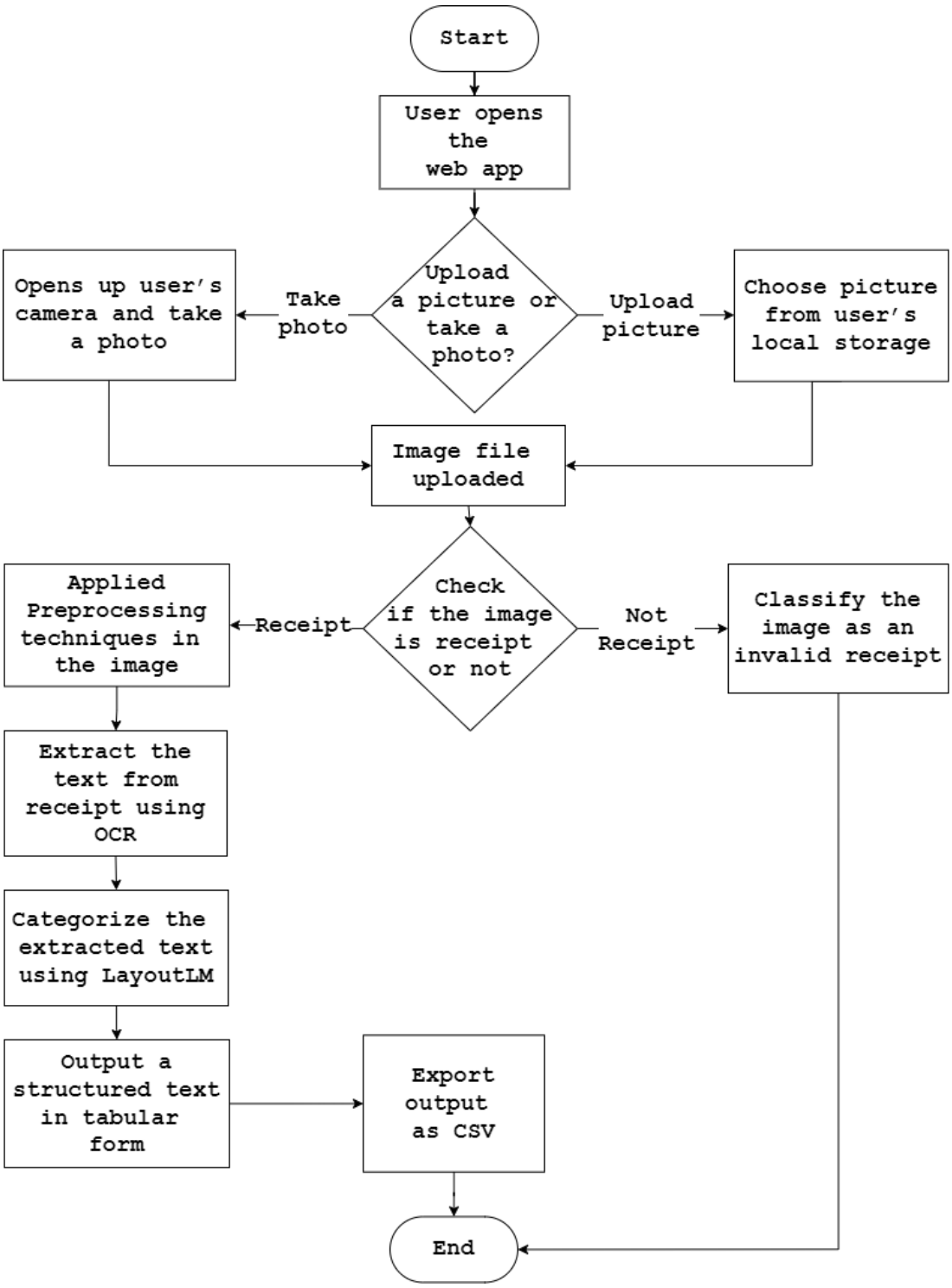



**Figure 5.** ExCeipt Web App Flowchart Diagram

As shown in Figure 5, it shows how the ExCeipt Web app works. The user is redirected to the web app display page and asks for the image of the receipt to be extracted. When the image is verified as a receipt, it will then proceed to extract the information from the receipt. After the extraction, it will ask the user the file type of the output.

Object-oriented design shows how a software system will be structured and organized around objects, classes, and their relationships to each other. It involves breaking down a software system into a set of modular, reusable objects that encapsulate data and behavior.

*Process Design*

The Process design outlines the necessary steps to carry out a task, operate a system, and ensure its implementation. Figure 6 depicts the process design for the system that has been proposed.

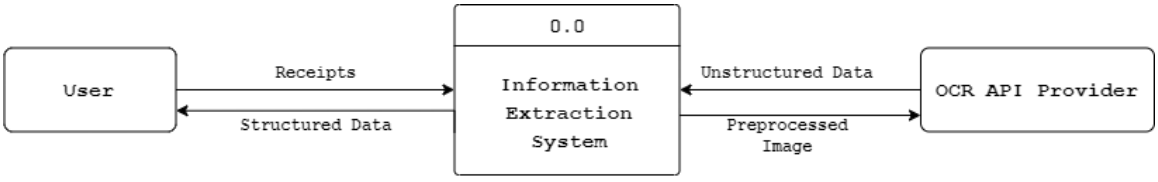



**Figure 6.** Context Diagram / Level 0 Diagram

At level 0 of the system, the user interacts with the system by providing data, receipts. The system extracts information from the receipts using an OCR API provider. Then, the extracted information is preprocessed into structured data.

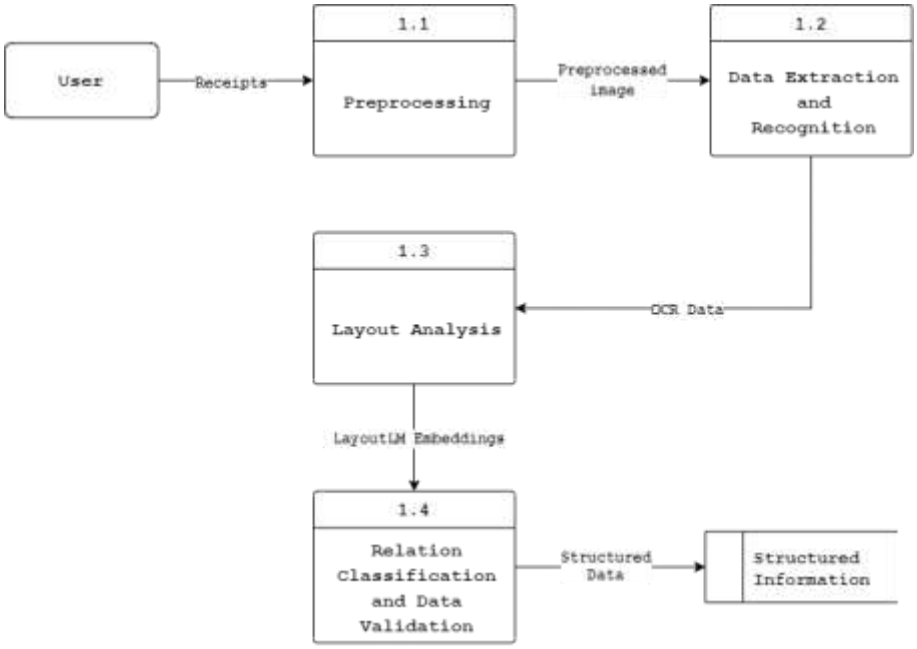



**Figure 7.** Level 1 Diagram

In level 1 of the diagram, the system preprocesses the image, extracts and recognizes data using OCR technology, and performs layout analysis. The resulting OCR data is processed and converted into numerical layout embeddings, which are validated and classified before being stored as structured

information. This information can be used for further analysis and information retrieval.

*System Development Life Cycle*

This study will utilize an Agile Software Development Life Cycle. Agile development is an iterative approach to software development that completes work in short increments called sprints (Microsoft learn, 2022). This kind of model allows for adaptability and quick response to change through its focus on collaborative decision-making, and development over multiple short cycles or sprints.

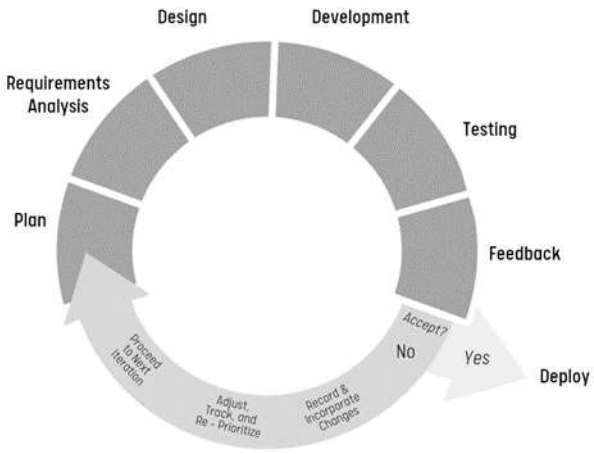



**Figure 8.** System Development Life Cycle Model

Figure 8 shows the seven phases for the System Development Life Cycle, namely: plan, requirements analysis, design, development, test, deploy, and feedback.

*Planning Phase*

During the first phase, the researchers define the scope and objectives of the thesis. They also create a high-level plan to achieve these goals, which includes identifying critical research questions, defining thesis goals, and creating a project roadmap. Furthermore, they choose appropriate research methods and set up the necessary tools and processes.

*Requirements Analysis*

In order to inform the study questions and hypotheses, the researchers collect and analyze data from a variety of sources throughout this stage. In order to understand the current state of knowledge on the subject, this phase involves conducting literature studies and interviews.

*Designing Phase*

They create a thorough research plan that includes selecting appropriate methods for data collection and analysis. Creating experiments, choosing suitable

algorithms or models for data analysis, and confirming the validity and dependability of the results are all included in this phase.

*Development Phase*

They carry out the study design by collecting and analyzing data during the development process in brief iterations or sprints. They deliver a potentially shippable chunk or chapter of their thesis, such as a draft chapter, at the end of each sprint.

*Testing Phase*

By using a variety of data triangulation techniques, the researchers evaluate and guarantee the accuracy of their findings during the testing phase. This includes contrasting our findings with those of currently used systems, performing more data analysis to support their conclusions, and getting feedback on their work from peers or advisors.

*Feedback*

Their major goal is to gather feedback from their advisers and peers in order to identify areas that might be improved in subsequent iterations. This entails asking for

input on a range of topics, including our writing, analysis, and findings presentation. They incorporate this criticism into their subsequent cycles of planning and development.

*Deployment*

The findings are presented to stakeholders during the deployment phase and/or publish them. They can submit their thesis for review or present their findings at conferences.

# CHAPTER 4 RESULTS AND DISCUSSION

## Implementation

The system was developed and tested on a Windows 11 Home operating system equipped with 8 GB of RAM and an Intel i5 12th Generation processor. The graphical processing needs were addressed using a combination of an NVidia GTX 1650 dedicated graphics card and the integrated Intel UHD Graphics.

For front-end development, a popular web development stack was employed, utilizing HTML, CSS, Javascript, and the JQuery library. The integrated development environment (IDE) of choice for front-end coding was Visual Studio Code.

The back-end development utilized Python 3.11.4. To provide a structured framework for the back-end, the Flask web application framework was implemented. Similar to the front-end, Visual Studio Code served as the primary coding environment. Notably, Kaggle Notebooks, a web-based platform for data science and machine learning, was leveraged for the specific task of fine-tuning and training the model.

Deployment of the application opted for a web-based service approach, eliminating the need for users to install any software locally. This enhances accessibility and simplifies the user experience.

*Data Collection*

Two distinct approaches were used to collect the dataset. First, the researchers collected their own personal receipts as well as receipts that participants freely gave. They also made contact with 7-Eleven, Iloilo Grace Pharmacy and WVSU Multi-Purpose Cooperative and requested additional receipts formally.

As soon as the researchers had physical copies of the receipts, they digitized them to make analysis easier. To produce digital copies, this required scanning each receipt with the appropriate tools, like a scanner. By digitizing the receipts, the researchers make sure that the data could be easily processed and analyzed using a certain computer-based method.


**Figure 9.** Gathered Dataset Sample

In Figure 9, we present a collection of sample receipts obtained from three distinct sources: 7-Eleven, Iloilo Grace Pharmacy, and WVSU Multi- Purpose Cooperative Merchants.

*Preprocessing*

There are some other processes or ways of preparing data for extracting and analyzing information by pre-processing taped transactions and acknowledgments and document receipts. This was done using online tools such as UBIAI Text Annotation tool for the annotation of images and Kaggle with GPU T4x2 accelerator for training the model.

*Dataset Annotation*


**Figure 10.** UBIAI Text Annotation Tool Interface

The researchers used a tool called UBIAI - Text Extraction to label texts from receipt images, as shown in Figure 10. They labeled 498 receipts in total, with 170 from 7-Eleven, 158 from Grace Pharmacy, and 170 from WVSU Multi-Purpose Coop.

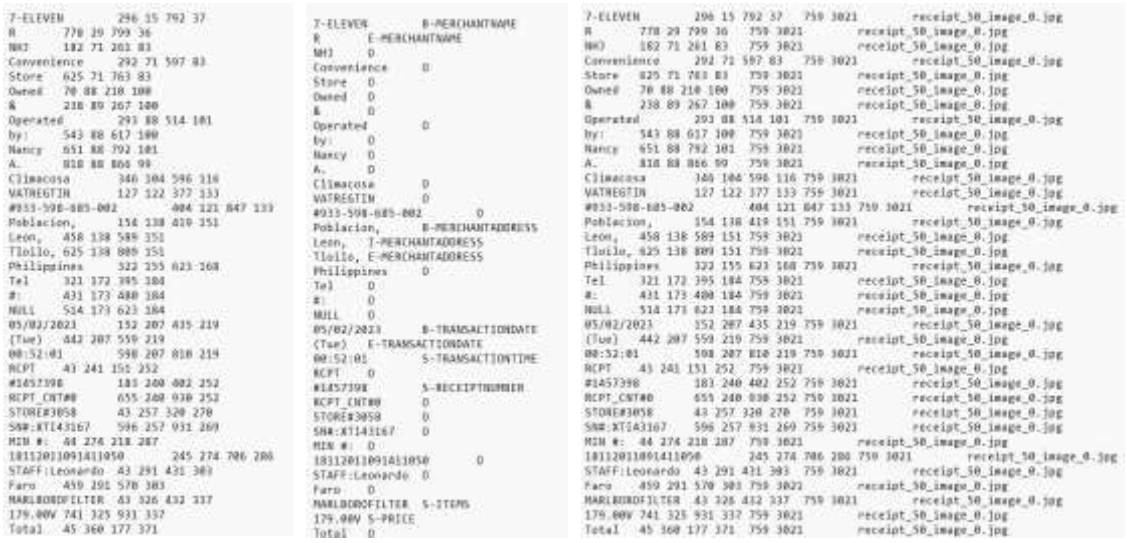


**Figure 11.** OCR Processed Output

From this work, they created a dataset with words labeled, their positions outlined, images, and a list of labels as shown in Figure 11. These files form an archive and will be used to fine-tune the LayoutLMv3 model. This refined dataset, made from careful annotation, will help improve the model's ability to understand and process receipt information accurately.

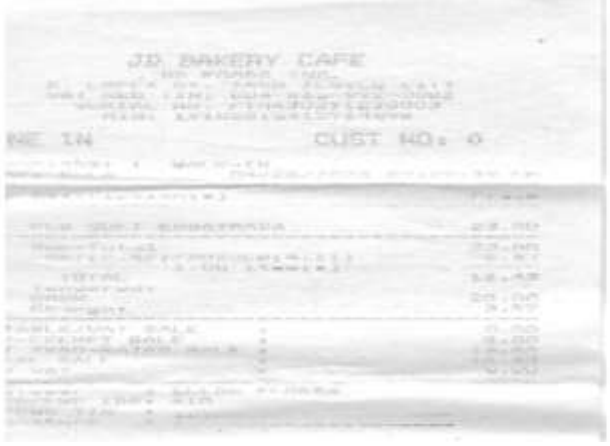
*Data Cleaning*

```
// Specify the folder path containing the images
folder_path <-
"/content/drive/MyDrive/THESIS/ExCeipt_Dataset/ExCeip
t-Dataset/Task1"

// Specify the blur threshold value
threshold <- 100

// Iterate over each file in the folder
for each filename in getFilesInFolder(folder_path) do
    image_path <- joinPaths(folder_path, filename)

    // Check if the path is a file
    if isFile(image_path) then
        // Read the image
        image <- readImage(image_path)

        // Check if the image is valid
        if image is not null then
            // Convert the image to grayscale
            gray <- convertToGrayscale(image)

            // Calculate the blur score using
Laplacian variance
            blur_score <-
calculateLaplacianVariance(gray)
```

**Figure 12.** Blurry Picture Detection Pseudocode

The researchers attempted to ensure the quality and integrity of the dataset by using a single criterion for assessing the image clarity. Removing or replacing blurry images helped to maintain the overall accuracy of the data, minimizing potential errors that could arise from unreadable or distorted content.

The image 'receipt_3.png' is blurry (Blur score: 73.32003642447437)




The image 'receipt_93.png' is blurry (Blur score: 88.63728619615082)

The image 'receipt_92.png' is blurry (Blur score: 51.806809686745986)

**Figure 13.** Blurry Picture Detection Results

Based on Figure 13, the blurry picture detection results show a clear distinction between sharp and blurred images.

*Data Normalization*

Receipts may include diverse representations of data, such as different date formats, currency symbols, or measurement units. Data normalization involves converting the extracted data into a standardized format, ensuring consistent representation across all receipts, such as normalizing the bounding box.

In this study,the researchers normalize the bounding box points within a range of [0, 1000]. The coordinates can be scaled using the width and height of the image. This rescaling ensures that the bounding box points are proportionally adjusted to fit within the desired range of values, facilitating consistent analysis and comparisons across different images.

*Image Enhancement*

```
def enhance_txt(img, intensity_increase=20, bilateral_filter_diameter=9, bilateral_filter_sigma_color=75,
bilateral_filter_sigma_space=75):
    # Get the width and height of the image
    w = img.shape[1]
    h = img.shape[0]
    w1 = int(w * 0.05)
    w2 = int(w * 0.95)
    h1 = int(h * 0.05)
    h2 = int(h * 0.95)
    ROI = img[h1:h2, w1:w2]  # 95% of the center of the image
    threshold = np.mean(ROI) * 0.88  # % of average brightness

    # Convert image to grayscale
    grayscale_img = cv2.cvtColor(img, cv2.COLOR_BGR2GRAY)

    # Apply Gaussian blur
    blurred = cv2.GaussianBlur(grayscale_img, (1, 1), 0)

    edged = 255 - cv2.Canny(blurred, 100, 150, apertureSize=7)

    # Increase intensity by adding a constant value
    img = np.clip(img + intensity_increase, 0, 255).astype(np.uint8)

    # Apply bilateral filter to reduce noise
    img = cv2.bilateralFilter(img, bilateral_filter_diameter, bilateral_filter_sigma_color, bilateral_filter_sigma_space)

    _, binary = cv2.threshold(blurred, threshold, 255, cv2.THRESH_BINARY)
    return binary
```

**Figure 14.** Image Enhancement

To enhance the process of extracting information from the receipt, the researchers have added some image enhancement techniques, such as calculating the Region of Interest, Threshold calculation, applying Gaussian Blur, increasing the intensity, and converting the image to binary.

*File Optimization*

The researchers identified a significant issue concerning the processing or scanning time for larger images or high-sized files, attributing this delay to their substantial size.

```
def preprocess_image(image_path, max_file_size_mb=1, target_file_size_mb=0.5):
    try:
        # Read the image
        image = cv2.imread(image_path)
        # Enhance text
        enhanced = enhance_txt(image)

        # Save the enhanced image to a temporary file
        temp_file_path = tempfile.NamedTemporaryFile(suffix='.jpg').name
        cv2.imwrite(temp_file_path, enhanced)

        # Check file size of the temporary file
        file_size_mb = os.path.getsize(temp_file_path) / (1024 * 1024)  # Convert to megabytes

        while file_size_mb > max_file_size_mb:
            print(f"File size ({file_size_mb} MB) exceeds the maximum allowed size ({max_file_size_mb} MB). Resizing the image.")
            ratio = np.sqrt(target_file_size_mb / file_size_mb)
            new_width = int(image.shape[1] * ratio)
            new_height = int(image.shape[0] * ratio)

            # Resize the image
            enhanced = cv2.resize(enhanced, (new_width, new_height))

            # Save the resized image to a temporary file
            temp_file_path = tempfile.NamedTemporaryFile(suffix='.jpg').name
            cv2.imwrite(temp_file_path, enhanced)

            # Update file size
            file_size_mb = os.path.getsize(temp_file_path) / (1024 * 1024)
            print(f"New file size: ({file_size_mb} MB)")

        # Return the final resized image
        image_resized = cv2.imread(temp_file_path)
        return image_resized

    except Exception as e:
        print(f"An error occurred in preprocess_image: {str(e)}")
        return None
```

**Figure 15.** File Optimization

To mitigate this problem, a practical solution was implemented. They incorporated a function specifically designed to address this concern. This function comes into play when an image surpasses the 1MB threshold; it automatically initiates a resizing process aimed at reducing the image's file size. By implementing this function, the researchers sought to optimize the processing time by shrinking the image size, thereby facilitating

faster and more efficient handling of larger images during scanning or processing tasks.

*Fine-Tuning the LayoutLMv3*

After preparing the dataset and implementing necessary preprocessing steps, the researchers proceeded by setting up the foundational components for fine-tuning the LayoutLMv3 model.

```
model = LayoutLMv3ForTokenClassification.from_pretrained("microsoft/layoutlmv3-base",
                                                         id2label=id2label,
                                                         label2id=label2id)

processor = AutoProcessor.from_pretrained("microsoft/layoutlmv3-base", apply_ocr=False)
```

**Figure 16.** LayoutLMv3 Base Model

Initially, they loaded the preprocessed dataset and established the crucial metrics and labels required for the subsequent training process. Alongside this, we acquired both the LayoutLMv3 model itself and its associated processor, essential components for the fine-tuning phase.

```
NUM_TRAIN_EPOCHS = 120
PER_DEVICE_TRAIN_BATCH_SIZE = 1
PER_DEVICE_EVAL_BATCH_SIZE = 1
LEARNING_RATE = 4e-5
```

```
training_args = TrainingArguments(output_dir="test",
                                  report_to="none",
                                  # max_steps=1500,
                                  num_train_epochs=NUM_TRAIN_EPOCHS,
                                  logging_strategy="epoch",
                                  save_total_limit=1,
                                  per_device_train_batch_size=PER_DEVICE_TRAIN_BATCH_SIZE,
                                  per_device_eval_batch_size=PER_DEVICE_EVAL_BATCH_SIZE,
                                  learning_rate=LEARNING_RATE,
                                  evaluation_strategy="no",
                                  save_strategy="no",
                                  # eval_steps=100,
                                  load_best_model_at_end=True,
                                  metric_for_best_model="f1")
```

```
# Initialize our Trainer
trainer = Trainer(
    model=model,
    args=training_args,
    train_dataset=train_dataset,
    eval_dataset=eval_dataset,
    tokenizer=processor,
    data_collator=default_data_collator,
    compute_metrics=compute_metrics,
)
```

**Figure 17.** Trainer and Training Arguments

With these foundational elements in place, we moved on to defining the training parameters and configurations necessary for the trainer. These configurations play a pivotal role in determining how the model will be trained, ensuring that it learns and adapts effectively to our specific use case.


**Figure 18.** Training Process

Upon completion of these preparatory steps, the model stands primed and ready for the fine-tuning phase. The convergence of dataset, metrics, labels, model architecture, processor, and training configurations forms the robust framework necessary for refining the LayoutLMv3 model through fine-tuning, allowing it to learn and improve its performance on the given task.

*Fine-Tuning CNN Resnet18 Model*

The researchers have utilized the fastai library for creating a convolutional neural network (CNN) learner object.

```
lr1 = 1e-3
lr2 = 1e-1
learn.fit_one_cycle(40,slice(lr1,lr2))
```

**Figure 19.** Training the CNN Model

In this study, the researchers utilized a pre- trained model, ResNet18 and trained this pre-trained model to recognize images of receipts and non- receipts.

*Text Extraction from Receipt Image*

Utilizing the OCR Space API, we've streamlined the extraction of text from receipt images, enhancing the precision and efficiency of our process. OCR Space offers a comprehensive solution for parsing both images and multi-page PDF documents, delivering extracted text results in a structured JSON format.

The workflow begins by submitting the receipt image to the OCR Space API, which swiftly processes it and returns a JSON file containing the extracted text, complete with bounding box coordinates for each word. This meticulous

output ensures the preservation of spatial information crucial for accurate text labeling.

To seamlessly integrate with their LayoutLMv3 model, the researchers conduct preprocessing on the JSON file, extracting only the words and their corresponding bounding boxes. These refined data are then forwarded to the model for text labeling, enabling precise identification and categorization of receipt elements.


```
"ParsedResults": [
  {
    "TextOverlay": {
      "Lines": [
        {
          "LineText": "¡7-Eleven®",
          "Words": [
            {
              "WordText": "7",
              "Left": 305,
              "Top": 227,
              "Height": 29,
              "Width": 123
            },
            {
              "WordText": "-",
              "Left": 305,
              "Top": 227,
              "Height": 29,
              "Width": 123
            },
            {
              "WordText": "Eleven",
              "Left": 432,
              "Top": 227,
              "Height": 29,
              "Width": 112
            },
            {
              "WordText": "®",
              "Left": 548,
              "Top": 227,
              "Height": 29,
              "Width": 154
            }
          ],
          "MaxHeight": 29,
          "MinTop": 227
        },
```


**Figure 20.** Sample Output of OCR Space Text Extraction

An example output of the texts extracted from a scanned 7-Eleven receipt.

*System Inputs and Outputs*

The following screenshots showcase the envisioned user interface of the application, representing the front-end design. These interfaces have been designed to provide a user-friendly and intuitive experience.


**Figure 21.** Upload Page

In this page, the user can upload his photos limited only to file format jpg and png. He can upload his photos from his local device and upload to the web app for further processing.


**Figure 22.** Receipt Extractor Page

The uploaded image file from the user will undergo the process of extraction. The text on the receipts are extracted into more structured data.


**Figure 23.** Extract Completion Page

In this page the user can now see the extracted data from his uploaded receipts. In this part, he can check and evaluate the extracted text. He can edit the extracted data

and download the extracted structured data into his local device for other purposes.

*Evaluation*

In the evaluation of the model, the researchers will utilize F1 Score metrics such as the Recall, Precision, and Loss.

Recall, also known as sensitivity or true positive rate, is a metric that measures the ability of a model to identify all positive instances correctly. It quantifies the proportion of actual positive instances that were correctly predicted as positive by the model. A high recall value indicates that the model is effective at identifying positive instances, reducing the likelihood of false negatives. False positives, which are negative occurrences that are mistakenly categorized as positive, are not taken into account.

Precision evaluates the model's ability to correctly recognize positive cases out of all instances labeled as positive. It is believed to give favorable outcomes when it is high, demonstrating that the model has a small percentage of false positives.

Instead of using the arithmetic mean, the F1 score, which combines precision and recall, is determined as the harmonic mean of these two measurements. While a low F1 score means that either precision or recall is low, a high F1 score indicates that both are high. The F1 score can also be weighted to give precision and recall at various levels of importance.

Loss is a fundamental indicator that is used to gauge a model's performance. It calculates the discrepancy between the intended output and the output the model projected. The loss offers information about the performance of the model by quantifying the total number of errors. A high loss shows that the model is underperforming because it is unable to fully capture the underlying patterns in the data. A low loss, on the other hand, indicates that the model is working well and accurately approaching the desired output. During model training, minimizing the loss is the main goal because it shows progress in the model's capacity for precise prediction. To improve model performance and reduce loss, a variety of strategies can be used.

## Results Interpretation and Analysis

*LayoutLMV3 Fine-Tuning Results*

The model underwent training using 498 receipts gathered explicitly for this study. To assess the model's performance accurately, several metrics were employed. These metrics included F1 Score, Precision, Recall, and Loss. They were utilized to evaluate and measure the model's accuracy, its ability to correctly identify and extract information, and its overall performance in handling receipt data.

**Table 1.**

*LayoutLMv3 Fine-Tuning Results on 7-Eleven Receipts*

| **Metric** | **Results** |
|---|---|
| Loss | 0.1680331975221634 |
| Precision | 0.9085545722713865 |
| Recall | 0.9112426035502958 |
| F1 Score | 0.9098966026587888 |
| Accuracy | 0.9856 |

| | |
|---|---|
| Runtime | 3.356 |
| Samples Per Second | 10.131 |
| Steps Per Second | 10.131 |
| Epoch | 120.0 |

**Table 2.**

*LayoutLMv3 Fine-Tuning Results on Iloilo Grace Pharmacy Receipts*

| Metric | Results |
|---|---|
| Loss | 0.04365146905183792 |
| Precision | 0.9683544303797469 |
| Recall | 0.9870967741935484 |
| F1 Score | 0.9776357827476039 |
| Accuracy | 0.9963985594237695 |
| Runtime | 3.2236 |
| Samples Per Second | 9.927 |

| | |
|---|---|
| Steps Per Second | 9.927 |
| Epoch | 120.0 |

**Table 3.**

*LayoutLMv3 Fine-Tuning Results on WVSU Multi-Purpose Coop Receipts*

| Metric | Results |
|---|---|
| Loss | 0.10417062789201736 |
| Precision | 0.9478672985781991 |
| Recall | 0.954653937947494 |
| F1 Score | 0.9512485136741974 |
| Accuracy | 0.9885117493472585 |
| Runtime | 3.3867 |
| Samples Per Second | 10.039 |
| Steps Per Second | 10.039 |
| Epoch | 120.0 |

Based on the results, the model trained on 7-Eleven receipts demonstrates a commendable accuracy of 0.986, albeit with a slightly higher evaluation loss of 0.168. However, its precision, recall, and F1 score, while solid, lag behind Grace Pharmacy and WVSU Coop. Grace stands out as the top performer, showcasing remarkable precision (0.968), recall (0.987), F1 score (0.978), and accuracy (0.996), along with the lowest evaluation loss of 0.044 among the three models. WVSU Coop follows closely with commendable precision (0.948), recall (0.955), and accuracy (0.989), positioning itself just behind Grace Pharmacy. While its evaluation loss is slightly higher than Grace's, Coop outperforms 7-Eleven across multiple metrics.

The F1 score result shows in this model a reasonably good balance between accurately identifying positive instances and minimizing false positives. The model can correctly categorize positive examples and minimize false positives, according to the Precision score. On the other hand, the model may miss some positive examples, according to the Recall score, even though it captures a sizable part of them. The model significantly reduces mistakes during training with a loss value.

These metrics suggest that the model performs reasonably well, achieving a good balance between precision and recall. However, there is still a need for adjustment and tweak to improve the model in capturing all positive instances.

After fine-tuning the model for 120 epochs, the following are the results of the evaluation as reflected in Table 4.

**Table 4.**

*Fine-Tuning LayoutLMv3 F1 Results*

| **Metric** | **Results** |
|---|---|
| Loss | 0.13155236840248108 |
| Precision | 0.9521988527724665 |
| Recall | 0.9549376797698945 |
| F1 Score | 0.9535662996649115 |
| Accuracy | 0.9912531676612442 |
| Runtime | 9.5208 |
| Samples Per Second | 10.503 |

| | |
|---|---|
| Steps Per Second | 10.503 |
| Epoch | 120.0 |

The results showed that the model has garnered a 99.12% accuracy, which is indicative of its exceptional capability to make precise predictions or classifications. Accompanied by a precision score of approximately 95.22% and a recall rate of 95.49%, the model demonstrates a robust ability to accurately identify relevant instances while minimizing false positives.

Moreover, achieving an evaluation loss of 0.1315 signifies the model's efficiency in learning and generalizing patterns from the data during its 120 epochs of training. This low loss value suggests that the model's predictions are close to the actual values, contributing to its high accuracy.


**Figure 24.** LayoutLMv3 Sample Inference results in 7-Eleven Receipt

The figure shows sample output from the LayoutLM Version 3 model, demonstrating its ability to extract and parse information from a 7-Eleven receipt image.

*CNN Model Results*

After training the pre-trained model with the receipt and non-receipt dataset, the researchers were able to attain the following results.

**Table 5.**

*CNN Model Training Results*

| Epoch | Train Loss | Valid Loss | Accuracy | Time |
|---|---|---|---|---|
| 1 | 0.229352 | 0.094690 | 0.975169 | 02:27 |
| 10 | 0.231085 | 0.002130 | 0.997743 | 02:27 |
| 20 | 0.172990 | 0.141639 | 0.988713 | 02:23 |
| 30 | 0.025840 | 0.105292 | 0.997743 | 02:35 |
| 40 | 0.025157 | 0.015434 | 0.997743 | 02:31 |

The training log showcases the model's progression across multiple epochs, revealing a consistent pattern of improvement in accuracy and validation loss. Starting with an initial accuracy of 97.52% and a training loss of 0.229352, the model steadily refined its performance. Across subsequent epochs, accuracy climbs to an impressive

99.77%, while the validation loss varies, showing occasional fluctuations. The training time remains relatively constant throughout, maintaining stability in efficiency. These results indicate the model's effective learning and its adeptness at extracting structured data from receipts. However, the slight variability in validation loss hints at potential areas for enhancing the model's generalization across different receipt samples. Overall, the model exhibits substantial progress and high accuracy, signifying its competence in processing and extracting relevant information from receipts.

The model has undergone a Classification matrix, which is a performance measurement tool for classification problems. The model has correctly identified 189 instances as non-receipts and 253 instances as receipts, which are known as True Negatives and True Positives respectively. On the other hand, the model incorrectly identified 0 instances as receipts when they were actually non-receipts, and 1 instance as a non-receipt when it was actually a receipt. Overall, the model seems to have a high accuracy in its classifications, as indicated by the high number of

True Positives and True Negatives and the low number of False Positives and False Negatives.

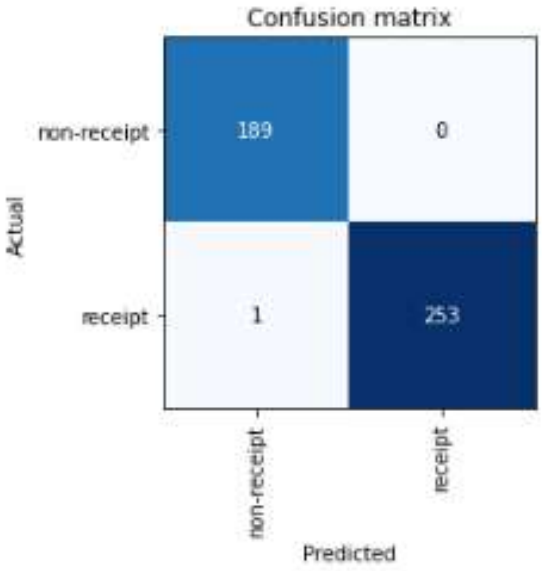



**Figure 25.** Confusion Matrix of the CNN Model

A confusion matrix shows a convolutional neural network's performance in classifying receipts and non-receipts, achieving high accuracy.

*System Evaluation Results*

The system was evaluated using ISO 25010 - Systems and software engineering - System and software quality requirements and evaluation (SQuaRE). System and software quality can be assessed through the use of this standard. The system will be evaluated in a variety of dimensions, such as Functional Stability, Performance Efficiency, Compatibility, Usability, Security, Reliability, Maintainability, Portability.

**Table 6.**

Summary of ISO/IEC 25010 Software Quality Standards Evaluation Result

| Characteristics | Mean | Description |
|---|---|---|
| Functional Completeness | 4.5 | Very Good |
| Performance Efficiency | 4.444 | Very Good |
| Compatibility | 4.5833 | Very Good |
| Usability | 4.5278 | Very Good |
| Security | 4.5 | Very Good |

| | | |
|---|---|---|
| Reliability | 4.41667 | Very Good |
| Maintainability | 4.46667 | Very Good |
| Portability | 4.56 | Very Good |
| OVERALL | 4.499 | Very Good |

Using the Likert Scale, the researchers gathered data for evaluation. Likert Scales are widely used tools that allow respondents to express their opinions. It consists of five answer options, ranging from strongly agree to strongly disagree, with a neutral option positioned in the middle. Both extreme poles are connected by intermediate answer options.

The evaluation was taken by 6 respondents, composed of Store Clerk, Accounting Professor and Accounting Students. The results of the evaluation, as presented in Figure 8, indicate that all eight characteristics assessed by the instrument, namely Functional Stability, Performance

Efficiency, Compatibility, Usability, Security, Reliability, Maintainability, and Portability are all "Very Good", with an overall mean of 4.39. This demonstrates the system's overall high level of quality and effectiveness in meeting the needs of users.

**Table 7.**

*ISO/IEC 25010 Software Quality Standards Evaluation Legend*

| Scales of Mean | Description |
|---|---|
| 5 - 4.1 | Very Good |
| 4 - 3.1 | Good |
| 3 - 2.1 | Fair |
| 2 - 1.1 | Poor |
| 1 | Very Poor |

# CHAPTER 5 SUMMARY, CONCLUSIONS, RECOMMENDATIONS

## *Summary of the Proposed Study Design and Implementation*

The model for receipt's data extraction, classification and detection was implemented in a user-friendly web application. The model is designed to take a scanned image of a receipt as input, subsequently validating the receipt's authenticity prior to engaging in text extraction and classification processes. Following validation, the model undergoes preprocessing and Optical Character Recognition (OCR) extraction. The resulting output is then fed into the model for the purpose of classifying and comprehending the text, assigning appropriate labels. Subsequently, the application returns an extracted text alongside its corresponding labels. Furthermore, a bounding box labeled image, indicating the precise location of the text, is generated, accompanied by its corresponding category.

## *Summary of Findings and Conclusions*

The results in receipt classification for validation garnered an accuracy score of 99.77%, and 99.12% for the text entity extraction and labeling, respectively.

LayoutLMv3 accumulated a very low evaluation loss of 0.13155, signifying the robustness and efficiency of the model in accurately classifying and extracting information from receipts. It also gained an accuracy of 99.12% and an F1 Score of 95.35%. This impressive performance underscores the effectiveness of LayoutLMv3 in handling diverse receipt formats and capturing key details with high precision.

Furthermore, the system underwent assessment through the ISO/IEC 25010 standard evaluation framework, showcasing a notable level of quality and efficacy in satisfying user requirements. In our system evaluation, all 8 characteristics examined received a “very good” rating, affirming the system's overall excellence in performance and user satisfaction.

*Recommendations*

To enhance results for future related studies, it is advisable to broaden the scope by including more merchants and incorporating additional datasets. Additionally, there is a need to extend the training duration to a larger epoch and integrate it as an additional step before the main training process. However, due to time constraints arising

from the manual annotation of receipts and the model's limitation to a specific format, a distinct platform was employed for manual annotation, incurring significant costs. Furthermore, the training of images is a resource-intensive task demanding substantial computing resources and time. Unfortunately, the researchers were unable to implement this approach, as it would have effectively doubled the overall training time.

Moreover, integrating the functionality to receive handwritten receipts would be a valuable enhancement to our web application. This addition could significantly expand the application's versatility and usability, allowing users to submit handwritten receipts, thus catering to a broader range of user preferences and needs.

# REFERENCES

Babu, S.C. (2022). *Digitize Receipts with Receipt OCR | Automated Receipt OCR*. Retrieved March 22, 2023, from https://nanonets.com/blog/receipt-ocr/

Brackert, T., Gossy, G., Haider, L., Keller, A., & Messenböck, R. (2021, October 13). *Simpler, faster, and more efficient operations in financial services. BCG Global*. Retrieved April 17, 2023, from https://www.bcg.com/publications/2019/simpler-faster-efficient-operations-financial-services

Bunch, E. (2021, November 16). *Information Extraction from Scanned Receipts: Fine-tuning LayoutLM on SROIE*. W&B. https://wandb.ai/wandb-data-science/layoutlm_sroie_demo/reports/Information-Extraction-from-Scanned-Receipts-Fine-tuning-LayoutLM-on-SROIE--VmlldzoxMjI5NzE2

Cabuz, A. (2019, June 25). *How to Use AI to Automate Invoices and Receipt Processing*. UiPath. Retrieved March 22, 2023, from https://www.uipath.com/blog/ai/ai-invoice-receipt-processing

Cabuz, A. (2019). *Using AI to Automate Invoices and Receipt Processing.* UiPath. Retrieved March 22, 2023, from https://www.uipath.com/blog/ai/ai-invoice-receipt-processing

Cambridge Dictionary (n.d.). *Automated*. Cambridge Dictionary. Retrieved March 22, 2023, from https://dictionary.cambridge.org/dictionary/english/automated

Conrad, S., CPA. (2023, March 26). *What is a Bookkeeper? - Definition | Meaning | Example*. My Accounting Course. Retrieved March 22, 2023, from https://www.myaccountingcourse.com/accounting-dictionary/bookkeeper

Contributor, T. (2022, November 15). *What is OCR (optical character recognition)?: Definition from TechTarget*. Retrieved April 17, 2023, from https://www.techtarget.com/searchcontentmanagement/definition/OCR-optical-character-recognition

Desk, S. (2023, March 6). *Accuracy in accounting and bookkeeping*. Vakilsearch. Retrieved March 27, 2023,

from https://vakilsearch.com/blog/importance-of-accuracy-in-accounting-and-bookkeeping/

Education, I. C. (2024, May 24). *What is optical character recognition (OCR)? - IBM blog*. IBM Blog. https://www.ibm.com/cloud/blog/optical-character-recognition

Fernandez, J., v-aangie., Bellarosa, C., Vivek, K., & Nert, J. (2023). *Receipt processing prebuilt AI model.* AI Builder | Microsoft Learn. Retrieved from March 22, 2023, https://learn.microsoft.com/en-us/ai-builder/prebuilt-receipt-processing

Fylehq. (2023). *Benefits of Automated receipt management for SMBs.* Retrieved March 27, 2023, from https://www.fylehq.com/blog/automated-receipt-management-for-smbs

Garncarek, Ł., Powalski, R., Stanisławek, T., Topolski, B., Halama, P., Turski, M., & Graliński, F. (2021). *LAMBERT: Layout-Aware Language Modeling for Information Extraction*. Retrieved March 26, 2023 from https://www.researchgate.net/publication/354357437_LAM

BERT_Layout-Aware_Language_Modeling_for_Information_Extraction

Hayes, A. (2022, December 27). *Receipt: Definition, types, and IRS rules*. Investopedia. Retrieved March 27, 2023, from https://www.investopedia.com/terms/r/receipt.asp

Huang, Y., Lv, T., Cui, L., Lu, Y., & Wei, F. (2022). *LayoutLMv3: Pre-training for Document AI with Unified Text and Image Masking*. Proceedings of the 30th ACM International Conference on Multimedia. https://doi.org/10.1145/3503161.3548112

Ionos. (n.d.). *Why are receipts important for good bookkeeping?* IONOS Startup Guide. Retrieved March 27, 2023, from https://www.ionos.com/startupguide/grow-your-business/good-bookkeeping-how-to-record-receipts-of-transactions/

Kithannnae, T. (2022, March 9). Layoutlm explained. Retrieved April 17, 2023, from https://nanonets.com/blog/layoutlm-explained/

Lajanuan. (2024). *Document Intelligence documentation - Quickstarts, Tutorials, API Reference - Azure AI*

*services*. Microsoft Learn. https://docs.microsoft.com/en-us/azure/applied-ai-services/form-recognizer/

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, *521*(7553), 436–444. https://doi.org/10.1038/nature14539

Liu, Y., Ott, M., Goyal, N., Du, J., Joshi, M., Chen, D., Levy, O., Lewis, M., Zettlemoyer, L., & Stoyanov, V. (2019, July 26). *Roberta: A robustly optimized Bert pretraining approach*. Retrieved March 26, 2023, from https://arxiv.org/abs/1907.11692

Lutkevich. (2020, January 1). What Is BERT (Language Model) and How Does It Work? Enterprise AI. Retrieved June 22, 2023, from https://www.techtarget.com/searchenterpriseai/definition/BERT-language-model

Makai, M. (n.d.). *Flask*. Full Stack Python. https://www.fullstackpython.com/flask.html

Mijacobs, V-thepet, JKirsch1, & EdKaim. (2022, November 29). *What is Agile Development? - Azure DevOps*. Microsoft Learn. Retrieved March 26, 2023, from

https://learn.microsoft.com/en-us/devops/plan/what-is-agile-development

Myers, L. (2021, May 4). *5 simple ways to automate bookkeeping*. Zapier. Retrieved March 22, 2023, from https://zapier.com/blog/simple-ways-to-automate-bookkeeping/

Nikolova, I. (2016, June 16). *Fundamentals: What is Information Extraction?* Ontotext. https://www.ontotext.com/knowledgehub/fundamentals/information-extraction/

P. on S. (2022, September 15). *What is accounting automation & why is it useful?* Retrieved April 17, 2023 from https://libeo.io/en/blog/accounting/accounting-automation?fbclid=IwAR0KLzE2fJFH7qo--xJ3iEFK30vdkiDfjNGhfDkAY4_Enz_OkHKpXY5FFvc

ProjectPro. (2022, November 24). *Bert NLP model explained for complete beginners*. ProjectPro. Retrieved March 26, 2023, from https://www.projectpro.io/article/bert-nlp-model-explained/558#

Rouse, M. (2011, August 18). *Encoding*. Techopedia. Retrieved March 22, 2023, from https://www.techopedia.com/definition/948/encoding

Staff, C. (2024, April 3). *What is Python used for? A beginner's guide*. Coursera. Retrieved April 17, 2023, from https://www.coursera.org/articles/what-is-python-used-for-a-beginners-guide-to-using-python

Tipalti (2023). *Accounting automation explained: How to get started*. Retrieved April 17, 2023, from https://tipalti.com/accounting-automation/

UBIAI. (2023, December 5). *UBIAI the Leader of Annotation tool*. Retrieved June 20, 2023, from https://ubiai.tools/ubiai-the-leader-of-annotation-tool-in-the-industry/

Vickneswaran, Anojan. (2016). *Overview on Accounting and Bookkeeping*. Retrieved from https://www.researchgate.net/publication/305681124_Overview_on_Accounting_and_Book_Keeping

X. Wang et al., *"LayoutLM: Pre-training of Text and Layout for Document Image Understanding,"* in Proceedings of

the 58th Annual Meeting of the Association for Computational Linguistics, 2020.

*What transformers can do*. What Transformers can do. (n.d.). Retrieved April 17, 2023, from https://huggingface.co/docs/transformers/task_summary#

Wood, T. (2020, July 14). *Convolutional neural network*. DeepAI. Retrieved April 17, 2023, from https://deepai.org/machine-learning-glossary-and-terms/convolutional-neural-network

Xu, Q., Wang, L., Liu, H., & Liu, N. (2022, January 1). *LayoutLM-Critic: Multimodal Language Model for Text Error Correction of Optical Character Recognition*. Communications in Computer and Information Science, 136–146. https://doi.org/10.1007/978-981-19-7943-9_11

Xu, Y. (2020, December 29). *LayoutLMv2: Multi-modal Pre-training for Visually-Rich Document Understanding*. arXiv.org. https://arxiv.org/abs/2012.14740

Xu, Y., Li, M., Cui, L., Huang, S., Wei, F., & Zhou, M. (2020). *LayoutLM: Pre-training of Text and Layout for Document Image Understanding*. Proceedings of the 26th

ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, 1192–1200. https://doi.org/10.1145/3394486.3403172

ZenBusiness Inc (2022). *How to Empower Your Work Using OCR (Guide for Accounting and Bookkeepers)*. Retrieved March 22, 2023, from https://www.zenbusiness.com/blog/how-to-empower-your-work-using-ocr-guide-for-accounting-and-bookkeepers/

# APPENDICES

# Appendix A - Letter to the Adviser

| | | | |
|---|---|---|---|
| | ADVISER'S ENDORSEMENT FORM (For Thesis Manuscript) | Document No. | WVSU-ICT-SOI-03-F10 |
| | | Issue No. | 1 |
| | | Revision No. | 0 |
| | WEST VISAYAS STATE UNIVERSITY | Date of Effectivity: | April 27, 2018 |
| | | Issued by: | CICT |
| | | Page No. | Page 1 of 1 |

Respectfully endorsed to the **Technical Editor**, the attached manuscript of the thesis entitled:

"Exceipt: Automated Bookkeeping Encoding Through Improved Information Extraction on Receipts"

Said manuscript has been presented to me for preliminary evaluation and guidance, and after a series of corrections/directions given which was implemented by the proponents whose names are listed hereunder and their thorough research, we have come to its completion.

Now therefore, I hereby **ENDORSE** the said thesis manuscript to the Technical Editor for **TECHNICAL EDITING**.

John Christopher Matco
Adviser's Name & Signature

Date: February 23, 2024

Group Members:
1. Dave Fagarita
2. Jimuel Servandil
3. Jannica Mae Magno
4. Jeziah Lois Catanus

**Note:** *This form should be accomplished and signed if the corrections and changes made by the adviser have been implemented and a new copy of the document have been printed for checking and submission to the next editor*

Appendix B - Letter of Request to the Technical Editor

| | TECHNICAL EDITOR'S ENDORSEMENT FORM (For Thesis Manuscript) | Document No. | WVSU-ICT-SOI-03-F11 |
|---|---|---|---|
| | | Issue No. | 1 |
| | | Revision No. | 0 |
| | WEST VISAYAS STATE UNIVERSITY | Date of Effectivity: | April 27, 2018 |
| | | Issued by: | CICT |
| | | Page No. | Page 1 of 1 |

**Respectfully endorsed to the English Editor, the attached manuscript of the thesis entitled:**

"Exceipt: Automated Bookkeeping Encoding Through Improved Information Extraction on Receipts"

**Said manuscript was presented to me and was reviewed and edited in terms of technical specifications, correctness of diagrams and other technical matters. The corrections and suggestions was carried and implemented by the proponents whose names are listed hereunder.**

**Now therefore, I hereby ENDORSE the said thesis manuscript to the English Editor/Grammarian for English Grammar Editing.**

Dr. Frank I. Elijorde
**Technical Editor's Name & Signature**

**Date:** March 27, 2024

**Group Members:**

1. Dave Fagarita
2. Jimuel Servandil
3. Jannica Mae Magno
4. Jeziah Lois Catanus

**Note:** ***This form should be accomplished and signed if the corrections and changes made by the Technical Editor have been implemented and a new copy of the document have been printed for checking and submission to the next editor.***

Appendix C - Letter of Request to the Grammarian

| | | | |
|---|---|---|---|
| | ENGLISH EDITOR/GRAMMARIAN'S ENDORSEMENT FORM (For Thesis Manuscript) | Document No. | WVSU-ICT-SOI-03-F12 |
| | | Issue No. | 1 |
| | | | |
| | | Revision No. | 0 |
| | WEST VISAYAS STATE UNIVERSITY | Date of Effectivity: | April 27, 2018 |
| | | Issued by: | CICT |
| | | Page No. | Page 1 of 1 |

Respectfully endorsed to the **Thesis Format Editor**, the attached manuscript of the thesis entitled:

"Exceipt: Automated Bookkeeping Encoding Through Improved Information Extraction on Receipts"

Said manuscript was presented to me for English grammar editing, corrections has been made and the proponents whose names are listed hereunder implemented said corrections and changes in the revised manuscript.

Now therefore, I hereby **ENDORSE** the said thesis manuscript for **Thesis Format Editing**.

Dr. Anapella Baga-An
English Editor/Grammarian's Name and Signature

Date: April 22, 2024

Group Members:
1. Dave Fagarita
2. Jimuel Servandil
3. Jannica Mae Magno
4. Jeziah Lois Catanus

*Note: This form should be accomplished and signed if the corrections and changes made by the English Editor have been implemented and a new copy of the document have been printed for checking and submission to the next editor.*

Appendix D - Letter of Request to the Format Editor

| | THESIS FORMAT EDITOR'S ENDORSEMENT FORM (For Thesis Manuscript) | Document No. | WVSU-ICT-SOI-03-F13 |
|---|---|---|---|
| | | Issue No. | 1 |
| | | | |
| | | Revision No. | 0 |
| | WEST VISAYAS STATE UNIVERSITY | Date of Effectivity: | April 27, 2018 |
| | | Issued by: | CICT |
| | | Page No. | Page 1 of 1 |

Respectfully endorsed to the **Thesis Coordinator**, the attached manuscript of the thesis entitled:

"Exceipt: Automated Bookkeeping Encoding Through Improved Information Extraction on Receipts"

Said manuscript was presented to me and has checked the preliminaries, thesis document convention and end matters, made some corrections which was implemented by the proponents whose names are listed hereunder.

Now therefore, I hereby **ENDORSE** said manuscript to the Thesis Coordinator for appropriate action.

Dr. Ma. Luche P. Sabayle
Thesis Format Editor's Name and Signature

Date: April 29, 2024

Group Members:
1. Dave Fagarita
2. Jimuel Servandil
3. Jannica Mae Magno
4. Jeziah Lois Catanus

Note: *This form should be accomplished and signed if the corrections and changes made by the Thesis Format Editor have been implemented and the four (4) new copies have been printed ready for bookbinding.*

Appendix E - Request Letter for Interview

**West Visayas State University**
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
Luna St., La Paz, Iloilo City 5000
Iloilo, Philippines
* Trunkline: (063) (033) 320-0870 to 78 loc. 1403 * Telefax No.: (033) 320-0879
* Website: www.wvsu.edu.ph * Email Address: cict@wvsu.edu.ph

March 29, 2023

**Basilia G. Galvan**
Head, Supply & Property Management Unit

Dear Ms. Galvan ,

We are writing to request your participation in a thesis study that we are conducting. We, the students of West Visayas State University College of Information and Communications Technology propose a study entitled "ExCeipt: Automated Bookkeeping Encoding through Improved Information Extraction on Receipts". We are reaching out to you because we believe that your insights and experiences would be invaluable to our research.

Our study aims to create an automated bookkeeping encoding system that improves information extraction from receipts. The purpose of this study is to gather information about your experiences with bookkeeping, receipt management, and any pain points you may have experienced.

The interview will take approximately 30 minutes and will be conducted in person. All information that you provide will be kept confidential and your identity will be protected.

Your participation in this study is voluntary, and you are free to withdraw at any time without consequence. Your participation would be greatly appreciated, and it will contribute to the success of this research project. We believe that your insights will be invaluable, and we would be honored if you could take the time to speak with us.

Respectfully yours,

**Jeziah Lois Catanus**
Researcher

**Jannica Mae Magno**
Researcher

**Dave Fagarita**
Researcher

**Jimuel Servandil**
Researcher

Noted:

Recommending Approval

**John Cristopher Mateo**
Research Instructor

**Dr. Ma. Beth S. Concepcion**
Dean, CICT

Appendix F - System Evaluation Questionnaire Form

# ExCeipt ISO/IEC Evaluation Questionnaire

Good Day. We are students from West Visayas State University - College of Information and Communications Technology, asking for your participation in our research entitled "ExCeipt: Automated Bookkeeping Encoding through improved Information Extraction on Receipts."

The web app that we created aims to automate the encoding process by automatically extracting information from receipt images, automatically label the information and display it in a structured and user-friendly way. All you have to do is upload an image of a receipt and it will give you the information that was extracted from the receipt.

**Instructions:**

1. Try the app by clicking this link: scezui-exceipt.hf.space

2. Upload your **7-11, Iloilo Grace Pharmacy or WVSU Coop Receipt**, or you can simply download one of the images below to try.

**NOTE**: As of now, the web app can only accurately extract information from 7-11, Iloilo Grace Pharmacy and WVSU Coop Receipts.

3. After Uploading and running the web app from the link provided, kindly evaluate the degree of compliance of the web app to the ISO/IEC 25010:2011 System and Software Quality Requirements and Evaluation Criteria by checking the column corresponding the degree to which you deemed the web app being evaluated complied or achieved using the scale below.

Thank you and enjoy.

**Images you can try:**

- Image 1
- Image 2
- Image 3

jeziahlois.catanus@wvsu.edu.ph Switch account

Not shared

Next Clear form

ExCeipt ISO/IEC Evaluation Questionnaire

jeziahlois.catanus@wvsu.edu.ph Switch account

Not shared

* Indicates required question

Untitled Section

Name (Optional)

Your answer

Position (Occupation) *

Your answer

E-mail Address (Optional)

Your answer

Back Next Clear form

# ExCeipt ISO/IEC Evaluation Questionnaire

jeziahlois.catanus@wvsu.edu.ph Switch account

Not shared

* Indicates required question

## Performance Efficiency

Performance relative to the amount of resources used under stated condition.

**Time Behavior.** Degree which response and processing times and throughput rates of a product or system, when performing its function, meet requirements. *

| | 1 | 2 | 3 | 4 | 5 | |
|---|---|---|---|---|---|---|
| Poor | ○ | ○ | ○ | ○ | ○ | Excellent |

**Resource Utilization.** Degree to which the amounts and types of resources used by a product or system, when performing its function, meet requirements. *

| | 1 | 2 | 3 | 4 | 5 | |
|---|---|---|---|---|---|---|
| Poor | ○ | ○ | ○ | ○ | ○ | Excellent |

**Capacity.** Degree to which the maximum limits of a product or system parameter meet requirements. *

| | 1 | 2 | 3 | 4 | 5 | |
|---|---|---|---|---|---|---|
| Poor | ○ | ○ | ○ | ○ | ○ | Excellent |

Back Next Clear form

# ExCeipt ISO/IEC Evaluation Questionnaire

jeziahlois.catanus@wvsu.edu.ph Switch account

Not shared

* Indicates required question

## Compatibility

Degree to which a product, system or component can exchange information with other products, systems are components, and/or perform its required functions, while sharing the same hardware or software environment.

**Co-existence**. Degree to which a product can perform its wired functions efficiently while sharing a common environment and resources with other products, without detrimental impact on any other product. *

| | 1 | 2 | 3 | 4 | 5 | |
|---|---|---|---|---|---|---|
| Poor | ○ | ○ | ○ | ○ | ○ | Excellent |

**Interoperability**. Degree to which two or more systems, products or components can exchange information and use information that has been exchange. *

| | 1 | 2 | 3 | 4 | 5 | |
|---|---|---|---|---|---|---|
| Poor | ○ | ○ | ○ | ○ | ○ | Excellent |

Back Next Clear form

## ExCeipt ISO/IEC Evaluation Questionnaire

jesiahlois.catamus@wvsu.edu.ph Switch account

Not shared

* Indicates required question

### Usability

Degree to which a product or system can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use.

**Appropriateness.** Recognizability. Degree to which users can recognize whether a product or system is appropriate for their needs. *

| | 1 | 2 | 3 | 4 | 5 | |
|---|---|---|---|---|---|---|
| Poor | ○ | ○ | ○ | ○ | ◉ | Excellent |

**Learnability.** Degree to which a product or system can be used by specified users to achieve specified goals of learning to use the product or system with effectiveness, efficiency freedom from risk and satisfaction in a specified context of use. *

| | 1 | 2 | 3 | 4 | 5 | |
|---|---|---|---|---|---|---|
| Poor | ○ | ○ | ○ | ○ | ◉ | Excellent |

**Operability.** Degree to which a product or system has attributes that make it easy to operate and control. *

| | 1 | 2 | 3 | 4 | 5 | |
|---|---|---|---|---|---|---|
| Poor | ○ | ○ | ○ | ○ | ◉ | Excellent |

**User-Error Protection.** Degree to which a system protects users against making errors. *

| | 1 | 2 | 3 | 4 | 5 | |
|---|---|---|---|---|---|---|
| Poor | ○ | ○ | ○ | ○ | ◉ | Excellent |

**User Interface Aesthetics.** Degree to which a user interface enables pleasing and satisfying interaction for the user. *

| | 1 | 2 | 3 | 4 | 5 | |
|---|---|---|---|---|---|---|
| Poor | ○ | ○ | ○ | ○ | ◉ | Excellent |

**Accessibility.** Degree to which a product or system can be used by people with the widest range of characteristics and capabilities to achieve a specified goal in a specified context of use. *

| | 1 | 2 | 3 | 4 | 5 | |
|---|---|---|---|---|---|---|
| Poor | ○ | ○ | ○ | ○ | ◉ | Excellent |

Back Next Clear form

# ExCeipt ISO/IEC Evaluation Questionnaire

jeziahlois.catanus@wvsu.edu.ph Switch account

Not shared

* Indicates required question

## Security

Degree to which a product or system protects information and data so that the persons or other products or systems have the degree of data access appropriate to their types and levels of authorization.

**Confidentiality.** Degree to which a product or system ensures the data are accessible only to those authorized to have access. *

| | 1 | 2 | 3 | 4 | 5 | |
|---|---|---|---|---|---|---|
| Poor | ○ | ○ | ○ | ○ | ○ | Excellent |

**Integrity.** Degree to which a system product or component prevents an authorized access to, or modification of, computer programs or data. *

| | 1 | 2 | 3 | 4 | 5 | |
|---|---|---|---|---|---|---|
| Poor | ○ | ○ | ○ | ○ | ○ | Excellent |

**Non-repudiation.** Degree to which actions or events can be proven to have taken place, so that the events are actions cannot be repudiated later. *

| | 1 | 2 | 3 | 4 | 5 | |
|---|---|---|---|---|---|---|
| Poor | ○ | ○ | ○ | ○ | ○ | Excellent |

**Accountability.** Degree to which the action of an entity can be traced uniquely to the entity. *

| | 1 | 2 | 3 | 4 | 5 | |
|---|---|---|---|---|---|---|
| Poor | ○ | ○ | ○ | ○ | ○ | Excellent |

Back Next Clear form

# ExCeipt ISO/IEC Evaluation Questionnaire

jezlahlois.catanus@wvsu.edu.ph Switch account

Not shared

* Indicates required question

## Reliability

Degree to which a system, product or component performs specified functions under specified conditions for specified period of time.

**Maturity.** Degree to which a system, product or component meets needs for reliability under normal operation. *

| | 1 | 2 | 3 | 4 | 5 | |
|---|---|---|---|---|---|---|
| Poor | ○ | ○ | ○ | ○ | ○ | Excellent |

**Availability.** Degree to which a system, product or component is operational and accessible when required for use. *

| | 1 | 2 | 3 | 4 | 5 | |
|---|---|---|---|---|---|---|
| Poor | ○ | ○ | ○ | ○ | ○ | Excellent |

**Fault Tolerance.** Degree to which a system, product or component operates as intended despite the presence of hardware or software faults. *

| | 1 | 2 | 3 | 4 | 5 | |
|---|---|---|---|---|---|---|
| Poor | ○ | ○ | ○ | ○ | ○ | Excellent |

**Recoverability.** Degree to which, in the event of an interruption or a failure, a product or system can recover the data directly affected and re-establish the desired state of the system. *

| | 1 | 2 | 3 | 4 | 5 | |
|---|---|---|---|---|---|---|
| Poor | ○ | ○ | ○ | ○ | ○ | Excellent |

Back Next Clear form

# ExCeipt ISO/IEC Evaluation Questionnaire

jestablois.catanuo@wvsu.edu.ph Switch account

Not shared

* Indicates required question

## Maintainability

Degree of effectiveness and efficiency with which a product or system can be modified by the intended maintainers.

**Modularity.** Degree to which a system or computer program is composed of discrete components such that a change to one component has minimal impact on other components. *

| | 1 | 2 | 3 | 4 | 5 | |
|---|---|---|---|---|---|---|
| Poor | ○ | ○ | ○ | ○ | ○ | Excellent |

**Reusability.** Degree to which an asset can be used in more than one system, or in building other assets. *

| | 1 | 2 | 3 | 4 | 5 | |
|---|---|---|---|---|---|---|
| Poor | ○ | ○ | ○ | ○ | ○ | Excellent |

**Analyzability.** Degree of effectiveness and efficiency with which it is possible to assess the impact on a product our system of an intended change to one or more of its parts, or to diagnose a product for deficiencies or causes of failures, or to identify parts to be modified. *

| | 1 | 2 | 3 | 4 | 5 | |
|---|---|---|---|---|---|---|
| Poor | ○ | ○ | ○ | ○ | ○ | Excellent |

**Modifiability.** Degree to which a product or system can be effectively and efficiency modified without introducing deffects or degrading existing product quality. *

| | 1 | 2 | 3 | 4 | 5 | |
|---|---|---|---|---|---|---|
| Poor | ○ | ○ | ○ | ○ | ○ | Excellent |

**Testability.** Degree of effectiveness and efficiency with which test criteria can be established for a system, product or component and test can be performed to determine whether those criteria have been met. *

| | 1 | 2 | 3 | 4 | 5 | |
|---|---|---|---|---|---|---|
| Poor | ○ | ○ | ○ | ○ | ○ | Excellent |

Back Next Clear form

# ExCeipt ISO/IEC Evaluation Questionnaire

jeziahloia.catanus@wvsu.edu.ph Switch account

Not shared

* Indicates required question

## Portability

Degree of effectiveness and efficiency with which a system, product or component can be transferred from one hardware, software or other operational or usage environment to another.

**Adaptability.** Degree to which a product or system can be effectively and efficiently be adapted for different or evolving hardware, software or operational or usage environment. *

| | 1 | 2 | 3 | 4 | 5 | |
|---|---|---|---|---|---|---|
| Poor | ○ | ○ | ○ | ○ | ○ | Excellent |

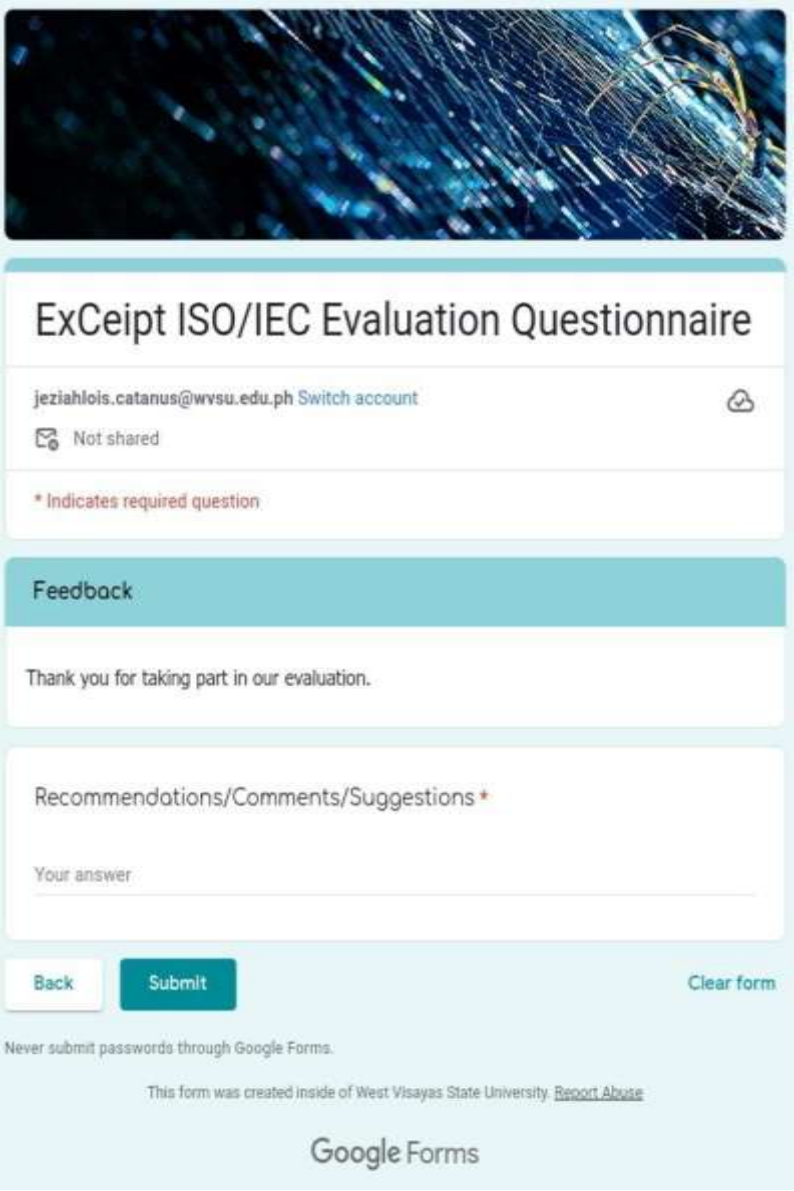
**Installability.** Degree of effectiveness and efficiency with which a product or system can be successfully installed and/or uninstalled in a specified environment. *

| | 1 | 2 | 3 | 4 | 5 | |
|---|---|---|---|---|---|---|
| Poor | ○ | ○ | ○ | ○ | ○ | Excellent |

**Replaceability.** Degree to which a product can replace another specified software product for the same purpose in the same environment. *

| | 1 | 2 | 3 | 4 | 5 | |
|---|---|---|---|---|---|---|
| Poor | ○ | ○ | ○ | ○ | ○ | Excellent |

Back Next Clear form


ExCeipt ISO/IEC Evaluation Questionnaire
jeziahlois.catanus@wvsu.edu.ph Switch account
Not shared
* Indicates required question
Feedback
Thank you for taking part in our evaluation.
Recommendations/Comments/Suggestions *
Your answer
Back
Submit
Clear form
Never submit passwords through Google Forms.
This form was created inside of West Visayas State University. Report Abuse
Google Forms

# Appendix G. Results: Training Data for LayoutLMv3

| Epoch | Step | Training Loss |
|---|---|---|
| 0 | 397 | 0.279800 |
| 1 | 794 | 0.062400 |
| 2 | 1191 | 0.053600 |
| 3 | 1588 | 0.047100 |
| 4 | 1985 | 0.044800 |
| 5 | 2382 | 0.041700 |
| 6 | 2779 | 0.040900 |
| 7 | 3176 | 0.038500 |
| 8 | 3573 | 0.037400 |
| 9 | 3970 | 0.037100 |
| 10 | 4367 | 0.033500 |
| 11 | 4764 | 0.033800 |
| 12 | 5161 | 0.031300 |
| 13 | 5558 | 0.028900 |
| 14 | 5955 | 0.026800 |
| 15 | 6352 | 0.025500 |
| 16 | 6749 | 0.022400 |
| 17 | 7146 | 0.021300 |
| 18 | 7543 | 0.019500 |
| 19 | 7940 | 0.015000 |
| 20 | 8337 | 0.014800 |
| 21 | 8734 | 0.013300 |
| 22 | 9131 | 0.010300 |
| 23 | 9528 | 0.011400 |
| 24 | 9925 | 0.009400 |
| 25 | 10322 | 0.007500 |
| 26 | 10719 | 0.006600 |
| 27 | 11116 | 0.006700 |
| 28 | 11513 | 0.005300 |
| 29 | 11910 | 0.004600 |
| 30 | 12307 | 0.005200 |
| 31 | 12704 | 0.004600 |
| 32 | 13101 | 0.002900 |
| 33 | 13498 | 0.004300 |
| 34 | 13895 | 0.004300 |
| 35 | 14292 | 0.006600 |
| 36 | 14689 | 0.004000 |
| 37 | 15086 | 0.003800 |
| 38 | 15483 | 0.006800 |
| 39 | 15880 | 0.003500 |
| 40 | 16277 | 0.003800 |
| 41 | 16674 | 0.002200 |
| 42 | 17071 | 0.003500 |
| 43 | 17468 | 0.001700 |
| 44 | 17865 | 0.002300 |
| 45 | 18262 | 0.001700 |
| 46 | 18659 | 0.002400 |
| 47 | 19056 | 0.002100 |
| 48 | 19453 | 0.002600 |
| 49 | 19850 | 0.003800 |
| 50 | 20247 | 0.001000 |
| 51 | 20644 | 0.001100 |

| | | |
|---|---|---|
| 52 | 21041 | 0.000700 |
| 53 | 21438 | 0.001100 |
| 54 | 21835 | 0.001300 |
| 55 | 22232 | 0.001700 |
| 56 | 22629 | 0.001400 |
| 57 | 23026 | 0.001700 |
| 58 | 23423 | 0.001400 |
| 59 | 23820 | 0.000300 |
| 60 | 24217 | 0.000600 |
| 61 | 24614 | 0.000500 |
| 62 | 25011 | 0.001800 |
| 63 | 25408 | 0.001300 |
| 64 | 25805 | 0.001000 |
| 65 | 26202 | 0.000500 |
| 66 | 26599 | 0.001000 |
| 67 | 26996 | 0.000800 |
| 68 | 27393 | 0.000700 |
| 69 | 27790 | 0.001100 |
| 70 | 28187 | 0.000900 |
| 71 | 28584 | 0.000600 |
| 72 | 28981 | 0.000500 |
| 73 | 29378 | 0.000400 |
| 74 | 29775 | 0.000400 |

| | | |
|---|---|---|
| 75 | 30172 | 0.001400 |
| 76 | 30569 | 0.000700 |
| 77 | 30966 | 0.002600 |
| 78 | 31363 | 0.000100 |
| 79 | 31760 | 0.001200 |
| 80 | 32157 | 0.000400 |
| 81 | 32554 | 0.000200 |
| 82 | 32951 | 0.000100 |
| 83 | 33348 | 0.000700 |
| 84 | 33745 | 0.000300 |
| 85 | 34142 | 0.000300 |
| 86 | 34539 | 0.000100 |
| 87 | 34936 | 0.000100 |
| 88 | 35333 | 0.000000 |
| 89 | 35730 | 0.000100 |
| 90 | 36127 | 0.000300 |
| 91 | 36524 | 0.000100 |
| 92 | 36921 | 0.000000 |
| 93 | 37318 | 0.000000 |
| 94 | 37715 | 0.000000 |
| 95 | 38112 | 0.000000 |
| 96 | 38509 | 0.000000 |
| 97 | 38906 | 0.000000 |

| | | |
|---|---|---|
| 98 | 39303 | 0.000400 |
| 99 | 39700 | 0.000000 |
| 100 | 40097 | 0.000100 |
| 101 | 40494 | 0.001300 |
| 102 | 40891 | 0.000000 |
| 103 | 41288 | 0.000000 |
| 104 | 41685 | 0.000000 |
| 105 | 42082 | 0.000000 |
| 106 | 42479 | 0.000000 |
| 107 | 42876 | 0.000000 |
| 108 | 43273 | 0.000000 |
| 109 | 43670 | 0.000000 |
| 110 | 44067 | 0.000000 |
| 111 | 44464 | 0.000000 |
| 112 | 44861 | 0.000300 |
| 113 | 45258 | 0.000000 |
| 114 | 45655 | 0.000000 |
| 115 | 46052 | 0.000000 |
| 116 | 46449 | 0.000000 |
| 117 | 46846 | 0.000000 |
| 118 | 47243 | 0.000000 |
| 119 | 47640 | 0.00000 |

# Appendix H. Results: Training Data for CNN

| epoch | train_loss | valid_loss | accuracy | time |
|---|---|---|---|---|
| 0 | 0.205229 | 0.059599 | 0.979684 | 02:25 |
| 1 | 0.142707 | 0.019781 | 0.993228 | 02:20 |
| 2 | 0.126255 | 0.027709 | 0.990971 | 02:20 |
| 3 | 0.102492 | 0.026335 | 0.997743 | 02:19 |
| 4 | 0.094030 | 0.226958 | 0.970655 | 02:22 |
| 5 | 0.171103 | 0.479346 | 0.981941 | 02:31 |
| 6 | 0.191729 | 0.399023 | 0.990971 | 02:18 |
| 7 | 0.289928 | 0.542277 | 0.972912 | 02:22 |
| 8 | 0.269095 | 0.198769 | 0.990971 | 02:22 |
| 9 | 0.299857 | 0.192198 | 0.995485 | 02:29 |
| 10 | 0.299107 | 0.312795 | 0.990971 | 02:22 |
| 11 | 0.368413 | 0.525500 | 0.988713 | 02:31 |
| 12 | 0.455837 | 0.282107 | 0.993228 | 02:22 |
| 13 | 0.519426 | 0.046707 | 0.997743 | 02:23 |
| 14 | 0.431784 | 0.062999 | 0.995485 | 02:20 |
| 15 | 0.333837 | 0.129406 | 0.993228 | 02:24 |
| 16 | 0.359816 | 0.066535 | 0.993228 | 02:22 |
| 17 | 0.313124 | 0.149032 | 0.995485 | 02:22 |
| 18 | 0.314941 | 0.166579 | 0.993228 | 02:23 |
| 19 | 0.446599 | 0.208862 | 0.990971 | 02:24 |
| 20 | 0.381422 | 0.076279 | 0.997743 | 02:22 |

| | | | | |
|---|---|---|---|---|
| 21 | 0.291001 | 0.106046 | 0.995485 | 02:22 |
| 22 | 0.204077 | 0.271010 | 0.988713 | 02:32 |
| 23 | 0.192935 | 0.000001 | 1.000000 | 02:20 |
| 24 | 0.153508 | 0.005572 | 0.997743 | 02:17 |
| 25 | 0.156020 | 0.000006 | 1.000000 | 02:22 |
| 26 | 0.135863 | 0.000000 | 1.000000 | 02:31 |
| 27 | 0.078742 | 0.002120 | 0.997743 | 02:20 |
| 28 | 0.081195 | 0.012585 | 0.997743 | 02:27 |
| 29 | 0.062399 | 0.055934 | 0.995485 | 02:25 |
| 30 | 0.083715 | 0.060890 | 0.997743 | 02:25 |
| 31 | 0.060076 | 0.000211 | 1.000000 | 02:25 |
| 32 | 0.036950 | 0.002544 | 0.997743 | 02:20 |
| 33 | 0.022070 | 0.011787 | 0.997743 | 02:26 |
| 34 | 0.016400 | 0.000052 | 1.000000 | 02:28 |
| 35 | 0.039930 | 0.000002 | 1.000000 | 02:23 |
| 36 | 0.027423 | 0.000001 | 1.000000 | 02:23 |
| 37 | 0.033601 | 0.000017 | 1.000000 | 02:22 |
| 38 | 0.025307 | 0.000009 | 1.000000 | 02:28 |
| 39 | 0.024325 | 0.000008 | 1.000000 | 02:25 |

# Appendix I - CNN Hyperparameter Tuning

| epoch | train_loss | valid_loss | accuracy | time |
|---|---|---|---|---|
| 0 | 0.028920 | 0.000124 | 1.000000 | 02:25 |
| 1 | 0.029602 | 0.039749 | 0.995485 | 02:27 |
| 2 | 0.047816 | 0.000000 | 1.000000 | 02:21 |
| 3 | 0.095908 | 0.040196 | 0.997743 | 02:26 |
| 4 | 0.117355 | 0.017502 | 0.997743 | 02:29 |
| 5 | 0.172963 | 0.022547 | 0.993228 | 02:25 |
| 6 | 0.162372 | 0.136571 | 0.993228 | 02:23 |
| 7 | 0.166488 | 0.062105 | 0.986456 | 02:23 |
| 8 | 0.159168 | 0.040274 | 0.995485 | 02:23 |
| 9 | 0.114997 | 0.016517 | 0.995485 | 02:28 |
| 10 | 0.093271 | 0.000019 | 1.000000 | 02:30 |
| 11 | 0.060940 | 0.000065 | 1.000000 | 02:25 |
| 12 | 0.068195 | 0.001144 | 1.000000 | 02:23 |
| 13 | 0.050627 | 0.001009 | 1.000000 | 02:25 |
| 14 | 0.034810 | 0.007881 | 0.997743 | 02:23 |
| 15 | 0.045892 | 0.000902 | 1.000000 | 02:25 |
| 16 | 0.036828 | 0.000000 | 1.000000 | 02:24 |
| 17 | 0.029192 | 0.000001 | 1.000000 | 02:19 |
| 18 | 0.025051 | 0.000008 | 1.000000 | 02:22 |
| 19 | 0.016698 | 0.000002 | 1.000000 | 02:26 |

# Appendix J - Training & Testing Dataset Screenshots

7/11 Receipts

WVSU Coop Receipts

Grace Pharmacy Receipts

Appendix K - Results: ISO/IEC Evaluation Questionnaire

| **Respondent** | **Position** |
|---|---|
| 1 | Accounting Student |
| 2 | Accounting Student |
| 3 | Store Clerk |
| 4 | Teacher |
| 5 | Teacher |
| 6 | Accounting Teacher |

| **Characteristics** | **Respondents** | | | | | | **Mean** | **Overall Mean** |
|---|---|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 | 6 | | |
| **Functionality Stability** | | | | | | | | |
| Functional Completion | 5 | 5 | 4 | 5 | 4 | 5 | 4.67 | 4.61 |
| Functional | 5 | 4 | 4 | 4 | 4 | 5 | 4.33 | |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Correctness | | | | | | | | |
| Functional Appropriateness | 5 | 5 | 5 | 5 | 4 | 5 | 4.83 | |
| **Performance Efficiency** | | | | | | | | |
| Time Behavior | 4 | 4 | 3 | 5 | 5 | 5 | 4.33 | 4.443 |
| Resource Utilization | 5 | 5 | 4 | 4 | 4 | 5 | 4.5 | |
| Capacity | 4 | 5 | 5 | 4 | 4 | 5 | 4.5 | |
| **Compatibility** | | | | | | | | |
| Co-existence | 5 | 5 | 4 | 4 | 4 | 5 | 4.5 | 4.585 |
| Interoperability | 5 | 5 | 4 | 4 | 5 | 5 | 4.67 | |
| **Usability** | | | | | | | | |
| Appropriateness | 5 | 5 | 4 | 4 | 5 | 5 | 4.67 | 4.53 |
| Learnability | 5 | 5 | 4 | 4 | 4 | 5 | 4.5 | |
| Operability | 5 | 4 | 4 | 4 | 5 | 5 | 4.5 | |
| User Error | 5 | 5 | 2 | 4 | 4 | 5 | 4.17 | |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Protection | | | | | | | | |
| User Interface Aesthetics | 5 | 5 | 5 | 4 | 4 | 5 | 4.67 | |
| Accessibility | 5 | 5 | 5 | 4 | 4 | 5 | 4.67 | |
| **Security** | | | | | | | | |
| Confidentiality | 5 | 5 | 3 | 5 | 4 | 5 | 4.5 | 4.5 |
| Integrity | 5 | 4 | 3 | 5 | 4 | 5 | 4.33 | |
| Non-repudiation | 5 | 5 | 4 | 4 | 5 | 5 | 4.67 | |
| Accountability | 5 | 5 | 3 | 4 | 5 | 5 | 4.5 | |
| **Reliability** | | | | | | | | |
| Maturity | 5 | 5 | 4 | 4 | 4 | 5 | 4.5 | 4.4175 |
| Availability | 5 | 4 | 4 | 4 | 5 | 5 | 4.5 | |
| Fault Tolerance | 4 | 5 | 3 | 4 | 4 | 5 | 4.17 | |
| Recoverability | 5 | 5 | 4 | 4 | 4 | 5 | 4.5 | |
| **Maintainability** | | | | | | | | |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Modularity | 5 | 5 | 4 | 4 | 3 | 5 | 4.33 | 4.466 |
| Reusability | 5 | 4 | 5 | 4 | 4 | 5 | 4.5 | |
| Analyzability | 5 | 5 | 3 | 4 | 4 | 5 | 4.33 | |
| Modifiability | 5 | 5 | 4 | 4 | 4 | 5 | 4.5 | |
| Testability | 5 | 5 | 4 | 4 | 5 | 5 | 4.67 | |
| **Portability** | | | | | | | | |
| Adaptability | 5 | 5 | 5 | 4 | 5 | 5 | 4.83 | 4.61 |
| Installability | 4 | 5 | 5 | 4 | 4 | 5 | 4.5 | |
| Replaceability | 5 | 5 | 4 | 4 | 4 | 5 | 4.5 | |
| **OVERALL** | | | | | | | | 4.2316875 |

# Appendix L - Disclaimer

## Disclaimer

This software project and its corresponding documentation entitled "ExCeipt: Automated Bookkeeping Encoding through Improved Information Extraction on Receipts" is submitted to the College of Information and Communications Technology, West Visayas State University, in partial fulfillment of the requirements for the degree, Bachelor of Science in Computer Science. It is the product of our own work, except for the utilization of the LayoutLMv3 model on HuggingFace Transformers.

We hereby grant the College of Information and Communications Technology permission to freely use, publish in local or international journal/conferences, reproduce, or distribute publicly the paper and electronic copies of this software project and its corresponding documentation in whole or in part, provided that we are acknowledged.

Dave F. Fagarita
Jimuel S. Servandil
Jannica Mae G. Magno
Jeziah Lois C. Catanus

June 2024