Internship Experience Report

CDA 571TUT INT1: Project Guidence

Company : Eitacies INC

Position : Machine Learning Engineer Intern

Project Name : CONFSEC

Reporting Manager : Kumar Ranjan Akash Dharshan(kdharshan@eitacies.com)

Duration : 30 May 2024 – 16 August 2024

Submitted to: Prof. Dr. Rachel Hageman Blair

15th August 2024

Indra Kiran Reddy Bonthu

University at Buffalo
Graduate School of Education

Introduction: During my internship at Eitacies Inc. as a Machine Learning Engineer Intern, I had the opportunity to work on a cutting-edge project aimed at enhancing the efficiency and security of conference calls applications such as Zoom, Webex, and MS Teams, through the development and implementation of advanced machine learning models. My primary focus was on video, audio and text classification tasks to improve detection systems, utilizing convolutional neural networks (CNNs) and various natural language processing (NLP) techniques. Currently, I am working on an emotion detection task using speech, employing the Toronto Emotional Speech Set and Long Short-Term Memory (LSTM) neural networks, a deep learning model. This experience not only allowed me to apply my existing skills but also challenged me to acquire new knowledge independently.

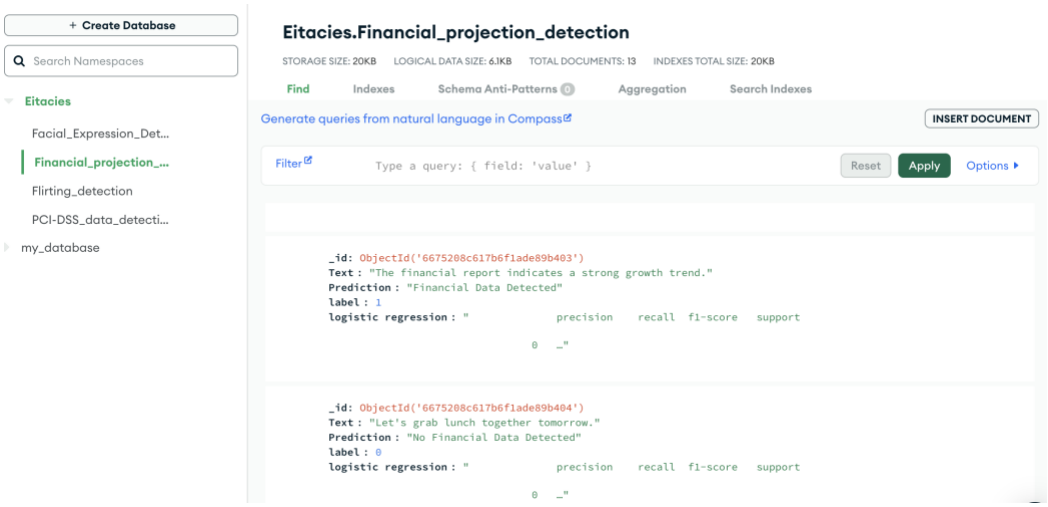
Work Summary: The core of my internship involved working on the development of advanced detection systems that could operate securely by preventing the sharing of prohibited content. The goal was to create datasets for text modules, create models that could accurately classify and detect various text, audio, and visual patterns during conference calls, ensuring a secure and seamless communication environment. The work required in-depth data analysis, model development, and constant iteration to improve accuracy and performance. The focus areas included identifying and flagging content related to PCI DSS (Payment Card Industry Data Security Standard), Financial Projection detection, Flirting Detection, human facial expressions using video and audio.

I contributed to several aspects of the project, including:

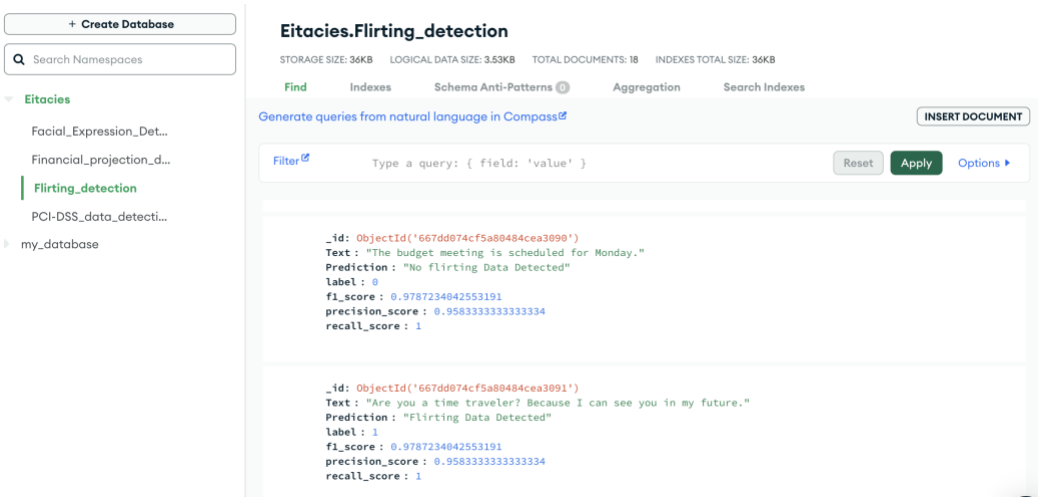
- **Data Analysis:** I've created data set on my own for the text classification tasks and performed extensive data preprocessing, feature extraction, and exploration using Python. This stage was crucial to ensure that the data fed into the models was of high quality and representative of the scenarios we aimed to classify.
- **Model Development:** Leveraging deep learning frameworks such as TensorFlow and PyTorch, I built and trained CNN models tailored to our specific needs. This involved selecting appropriate architectures, tuning hyperparameters, and implementing techniques to prevent overfitting and ensure robustness.
- **NLP Integration:** For text classification, I explored natural language processing (NLP) techniques to enhance the model's ability to understand context, detect sentiment, and identify specific keywords or phrases that might indicate security risks during conference calls.

Task Descriptions:

1. **Financial Projection Detection:**
    - My initial task was developing a model to detect financial projections during conference calls. The aim was to identify discussions related to future financial estimates or sensitive business forecasts to prevent accidental disclosure. In this task, I applied a combination of Bag of Words, TF-IDF. Additionally, I incorporated Bag of N-grams to capture context-specific phrases often used in financial discussions. This approach enabled the model to recognize and categorize financial projection-related content with higher precision.
    - **Output:** The model was integrated into our security system, providing real-time alerts when sensitive financial information was detected, thereby allowing for prompt intervention to prevent potential leaks. The flagged data was converted into JSON format and stored in MongoDB, ensuring seamless integration with the backend system for further analysis and action.



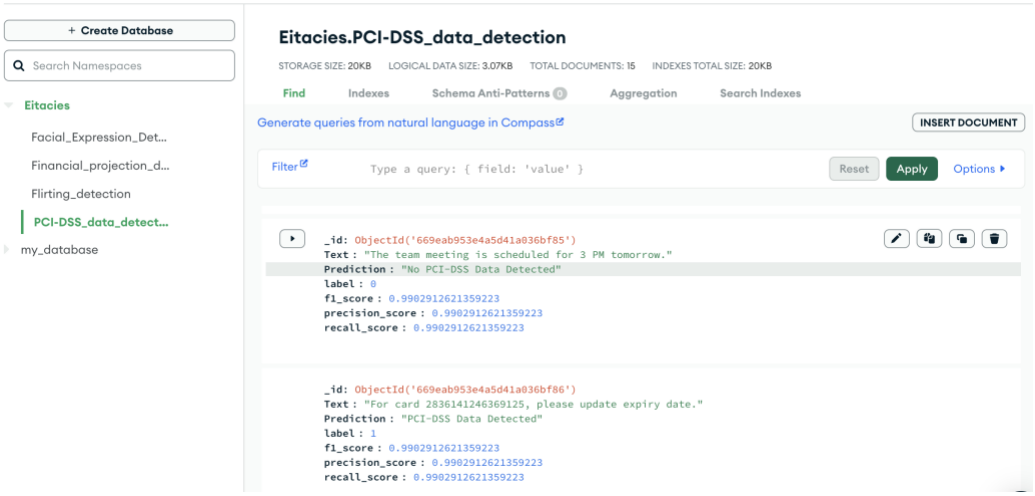
2. **Flirting Detection:**
   - In this task worked on detecting inappropriate behavior, such as flirting, during professional conference calls. This involved analyzing text data to identify language patterns indicative of unprofessional behavior. I utilized NLP techniques such as Bag of Words and TF-IDF to preprocess and vectorize the text data. The model was trained on a curated dataset and integrated into the broader detection system to monitor and flag inappropriate behavior in real-time.
   - **Output:** The resulting model was integrated into the broader detection system, enabling real-time monitoring, and flagging of inappropriate behavior during conference calls. The flagged instances were converted into JSON format and stored in MongoDB, ensuring efficient tracking and reporting.



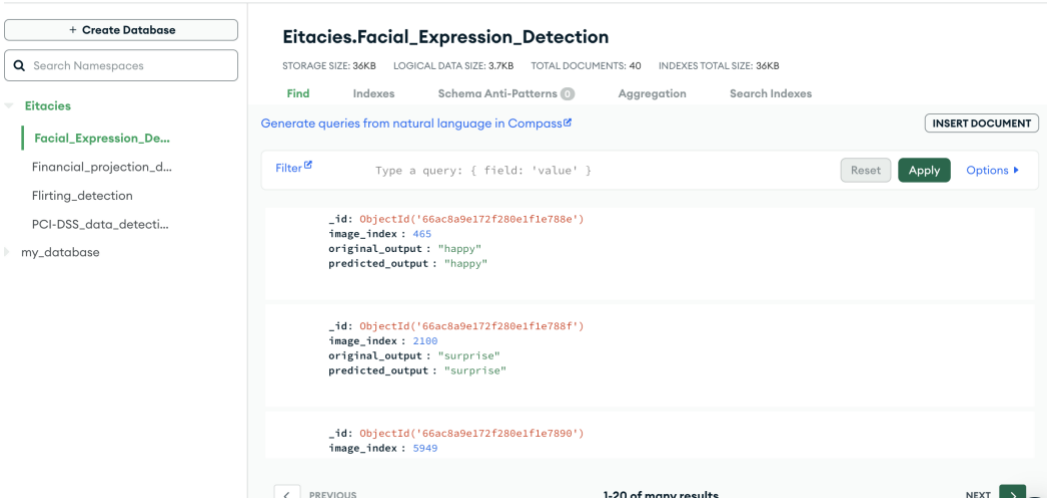
3. **PCI-DSS Detection:**
   - Ensuring compliance with Payment Card Industry Data Security Standards (PCI-DSS) was another critical task. My responsibility was to develop a system that could detect and flag any potential violations of PCI-DSS during conference calls. For this, I employed various NLP techniques to analyze the text transcriptions of spoken dialogue. I used methods like Term Frequency-Inverse Document Frequency (TF-IDF) and Bag of Words to extract meaningful features from the text data. Additionally, I explored Word2Vec to capture

semantic relationships between words, which helped in identifying subtle violations. The resulting model was capable of real-time analysis, scanning transcriptions for keywords and patterns that could indicate non-compliance, and alerting the system for immediate action.

- **Output:** The model was capable of real-time analysis, scanning transcriptions for keywords and patterns that could indicate non-compliance. The flagged data was converted into JSON format and stored the data predicted from best performed NLP techinique in MongoDB, allowing for immediate alerts and seamless integration with the backend system for compliance monitoring.



4. **Facial Expression Detection:**
   - Developing a model to detect facial expressions during video conferences. The objective was to enhance the system's ability to recognize and interpret participants' emotional states, which is crucial for maintaining security and effective communication in a conference setting. I utilized CNNs to process and classify images of facial expressions. To improve the model's accuracy, I implemented data augmentation techniques, which helped to diversify the training data. I experimented with various CNN architectures, to find the best fit for our needs.
   - **Output:** The output of this model was integrated into our detection system, enabling real-time analysis of video feeds to monitor participants' emotions. The detected emotions were converted into JSON format with detailed context and stored in MongoDB for seamless backend integration, facilitating the prompt identification and flagging of any unusual behavior.



5. **Current Work - Emotion Detection Using Speech:**
   - I am currently working on an emotion detection task using speech data, which is crucial for understanding the emotional context in conference calls. For this task, I am using the Toronto Emotional Speech Set, a well-known dataset that contains recordings of speech with various emotional tones. To process and classify the emotional content of speech, I am employing Long Short-Term Memory (LSTM) neural networks, which are well-suited for sequential data like audio. The LSTM model is designed to capture the temporal dependencies in speech, allowing it to better recognize the emotional patterns.
   - **Output:** I've completed training and validating the model and was left with converting the data into JSON format and storing it into MongoDB

	actualvalues	predictedvalues
20	disgust	disgust
21	surprise	surprise
22	sad	sad
23	neutral	neutral
24	disgust	disgust
25	surprise	surprise
26	disgust	disgust
27	neutral	neutral
28	angry	angry

This detailed breakdown covers the specific methodologies and approaches used for each task, reflecting the depth and breadth of work involved in enhancing the security and efficiency of conference call applications.

Skills Used and Learned:

- **Python Programming:** I extensively used Python for scripting, data manipulation, and implementing machine learning models. Libraries like Pandas, NumPy, Scikit-learn, and TensorFlow were integral to my work.
- **Machine Learning:** I applied various machine learning techniques, including Logistic Regression, Decision Trees, SVM, CNNs, and LSTM, focusing on fine-tuning models for better performance.
- **Deep Learning and CNNs:** My work required a solid understanding of CNNs, particularly in tasks involving image and video classification. I explored various architectures and implemented them to enhance model performance.
- **NLP Techniques:** I applied several NLP methods, including TF-IDF, Bag of Words, Word2Vec, and Bag of N-grams, to extract and analyze textual data. This was crucial in developing models that could understand and classify text-based information accurately.
- **LSTM Neural Networks:** For the emotion detection task, I am leveraging LSTM neural networks to analyze and classify speech data, allowing for the detection of emotions based on audio input.
- **Data Analysis:** I performed thorough data analysis using tools like Pandas and Matplotlib to ensure that the data was well-prepared for model training.
- **Independent Learning:** The internship required me to learn independently, especially in areas like advanced NLP techniques, Deep Learning techniques and model optimization. I regularly consulted technical documentation, and online resources to expand my knowledge.

Engagement Level: Throughout the internship, I maintained a high level of engagement with both my team and fellow interns. Regular collaboration and communication were key to the project's success. I was actively involved in team meetings, where we discussed progress, shared ideas, and provided feedback on every task.

My supervisor provided consistent guidance and mentorship, helping me navigate challenges and offering insights into best practices. This support was instrumental in helping me refine my skills and approach to problem-solving.

General Experience: The internship at Eitacies Inc. was a transformative experience that provided me with hands-on experience in solving real-world problems using machine learning, NLP and Deep Learning techniques. It deepened my understanding of the complexities involved in developing secure and efficient detection systems and allowed me to apply my academic knowledge in a professional context. The collaborative environment and technical challenges made this internship a significant milestone in my career development.

While I am bound by confidentiality constraints that limit the specifics I can share, the skills and experience I gained during this internship have been invaluable. They have equipped me with the tools necessary to excel in my future endeavors in the field of machine learning.

Conclusion: My time at Eitacies Inc. as Machine Learning Engineer has been a highly rewarding experience, enabling me to apply my skills, learn independently, and collaborate with a talented team. The knowledge and experience gained during this period will be instrumental in my continued growth in the field of machine learning, and I am excited to apply these skills in future projects.