Department of Technical Education

457, KLS’s Shri Vasantrao Potdar Polytechnic, Tilakwadi, Belagavi.

Dept. of Computer Science and Engg.

FORMAT 7

Weekly Status Report

Capstone project Name: **PhishAlert AI**

**Capstone project Members:**

* Aditya W
* Abhishek H
* Muzammil S

**Status: The PhishAlert AI project is currently in progress. The team has completed the initial setup phase and is now actively engaged in data collection and preprocessing. Model selection and training are scheduled to begin in the upcoming week.**

**Progress:**

**Week 1: Project Kickoff and Initial Setup: The project kick-off meeting was conducted, roles were assigned, and initial discussions on phishing threats and AI/ML's role in mitigation were held. Communication channels and project management tools were set up. A literature review and tool evaluation were completed, and data collection planning commenced.**

**Week 2: Data Collection and Preprocessing: Data collection is ongoing, with significant progress made in dataset size. Preliminary data analysis and feature identification have started. Initial steps for handling data imbalance and setting up the AI development environment have been completed.**

**Week 3: Model Selection and Training: Model selection has begun, with several models identified for experimentation. Feature refinement and model retraining are in progress. Work on developing a validation framework and implementing cross-validation techniques is underway.**

**Week 4: Model Optimization: In-depth performance evaluation of models has been conducted, leading to the identification of promising neural network-based models. Optimization of the neural network architecture and hyperparameter tuning are ongoing. The focus has also shifted toward enhancing the user interface for model deployment.**

**Week 5: Final Model Training and Initial Deployment: The final training of the optimized model is underway, with preparations for deployment in a simulated environment. Monitoring of model training progress and evaluation in a real-world scenario has commenced.**

**Highlights:**

* **Successful completion of project kick-off and initial setup.**
* **Significant progress in data collection and preprocessing.**
* **Identification of promising models for phishing detection.**
* **Ongoing optimization of model performance and user interface design.**

**Risks or Issues List:**

| **Risk/Issue** | **Status** | **Resolution** |
| --- | --- | --- |
| **Data Collection Challenges** | **Completed** | **N/A** |
| **Model Overfitting Concerns** | **Ongoing** | **N/A** |
| **User Interface Complexity** | **New** | **N/A** |

**Week 1: Project Kickoff and Initial SetupP**

**Tasks in Progress or Completed:**

| **Task Name** | **Description** | **Team Member Responsible** | **Percentage Complete** |
| --- | --- | --- | --- |
| Kick-off Meeting | Conducted a meeting to outline project scope, assign roles, and discuss initial strategies. | Abhishek H | 100% |
| Literature Review | Conducted a deep dive into the literature on phishing detection using AI/ML. Identified key sources. | Aditya W | 100% |
| Tool Evaluation | Evaluated and selected project management tools and AI/ML development environments. | Muzammil S | 100% |
| Communication Setup | Set up communication channels and project management tools for effective collaboration. | Abhishek H | 100% |
| Data Collection Plan | Developed a detailed plan for data collection, including sourcing phishing and benign emails. | Abhishek H, Aditya W | 100% |
| Data Collection | Started collecting data, focusing on publicly available datasets and reaching out to industry partners for proprietary data. | Muzammil S, Aditya W | 50% |

**Upcoming Tasks:**

| **Task Name** | **Description** | **Team Member Responsible** |
| --- | --- | --- |
| Data Collection | Continue collecting data and expanding the dataset. | Abhishek H |
| Data Analysis | Begin preliminary data analysis to understand dataset characteristics. | Muzammil S |
| Feature Identification | Identify features for phishing email detection. | Muzammil S, Aditya W |
| Data Imbalance Handling | Explore techniques for handling imbalanced datasets. | Abhishek H |
| AI Environment Setup | Set up the AI development environment for model training. | Muzammil S, Aditya W |
| Model Training Preparation | Prepare for model training by defining selection criteria and initial setup. | Abhishek H |

**Week 2: Data Collection and Preprocessing**

**Tasks in Progress or Completed:**

| **Task Name** | **Description** | **Team Member Responsible** | **Percentage Complete** |
| --- | --- | --- | --- |
| Data Collection | Continued with data collection, achieving a significant milestone in dataset size. | Muzammil S | 100% |
| Data Analysis | Began preliminary data analysis to understand dataset characteristics. | Abhishek H | 50% |
| Feature Identification | Identified a range of features for phishing email detection, including URL analysis and text pattern recognition. | Muzammil S, Aditya W | 70% |
| Data Imbalance Handling | Explored techniques for handling imbalanced datasets and implemented data augmentation methods. | Abhishek H | 80% |
| AI Environment Setup | Set up the AI development environment and prepared for model training. | Muzammil S, Aditya W | 100% |
| Model Training Preparation | Conducted a session on model selection criteria and prepared for model training. | Aditya W | 100% |

**Upcoming Tasks:**

| **Task Name** | **Description** | **Team Member Responsible** |
| --- | --- | --- |
| Model Training | Begin training models on the full dataset. | Muzammil S |
| Feature Extraction | Start feature extraction from the collected dataset. | Muzammil S |
| Data Cleaning Optimization | Optimize data cleaning scripts and procedures for efficiency. | Abhishek H , Aditya W |
| Model Evaluation Setup | Prepare for model evaluation by defining validation metrics and criteria. | Muzammil S, Aditya W |
| Documentation | Start documenting the data collection and preprocessing processes. | Abhishek H |

**Week 3: Model Selection and Training**

**Tasks in Progress or Completed:**

| **Task Name** | **Description** | **Team Member Responsible** | **Percentage Complete** |
| --- | --- | --- | --- |
| Model Selection | Selected several models for experimentation, including SVM, Random Forest, and simple Neural Networks. | Abhishek H | 100% |
| Feature Refinement | Refined features based on model feedback, focusing on enhancing text analysis capabilities using NLP techniques. | Muzammil S | 80% |
| Model Retraining | Retrained models with refined features and saw a notable improvement in model accuracy. | Abhishek H, Aditya W | 90% |
| Validation Framework | I began developing a validation framework to assess model performance using unseen data. | Muzammil S, Aditya W | 60% |
| Cross-Validation | Implemented cross-validation techniques to ensure model robustness and generalizability. | Muzammil S | 50% |

**Upcoming Tasks:**

| **Task Name** | **Description** | **Team Member Responsible** |
| --- | --- | --- |
| Model Evaluation | Continue with an in-depth evaluation of trained models to assess their performance. | Muzammil S |
| Hyperparameter Tuning | Begin optimizing model parameters using grid search and other tuning techniques. | Abhishek H , Aditya W |
| User Interface Design | Start designing the user interface for the model's deployment, focusing on real-time phishing detection alerts. | Muzammil S |
| Documentation | Proceed with documenting the model selection and training process for future reference. | Abhishek H , Aditya W |
| Team Meeting | Conduct a meeting to review progress and discuss any challenges encountered. | Aditya W |

**Week 4: Model Optimization**

**Tasks in Progress or Completed:**

| **Task Name** | **Description** | **Team Member Responsible** | **Percentage Complete** |
| --- | --- | --- | --- |
| Model Performance Evaluation | Conducted in-depth performance evaluation of all models and identified neural network-based models showing promising results. | Muzammil S | 100% |
| Neural Network Optimization | Began optimizing the neural network architecture by experimenting with different layers and activation functions. | Abhishek H | 50% |
| Hyperparameter Tuning | Implemented hyperparameter tuning using grid search techniques to improve model accuracy. | Abhishek H | 70% |
| User Interface Enhancement | Focused on improving the user interface for the model's deployment by designing a dashboard for real-time phishing detection alerts. | Muzammil S | 30% |
| Team Review | Conducted a team review of the current model's performance and user interface, identifying areas for improvement. | Aditya W | 100% |

**Upcoming Tasks:**

| **Task Name** | **Description** | **Team Member Responsible** |
| --- | --- | --- |
| Model Training Continuation | Continue model training with optimized parameters on the full dataset. | Abhishek H |
| UI Development | Progress with the development of the user interface for improved usability and functionality. | Muzammil S |
| Documentation Update | Update documentation to reflect changes made during model optimization and user interface enhancement. | Abhishek H |
| Stakeholder Meeting | Schedule and conduct a meeting with stakeholders to present progress and gather feedback on recent developments. | Aditya W |

**Week 5: Final Model Training and Initial Deployment**

**Tasks in Progress or Completed:**

| **Task Name** | **Description** | **Team Member Responsible** | **Percentage Complete** |
| --- | --- | --- | --- |
| Final Model Training | Kicked off the final training of the optimized model on the enhanced dataset. | Muzammil S | 100% |
| Model Training Monitoring | Monitored model training progress to ensure optimal resource utilization. | Muzammil S | 100% |
| Simulation Analysis | Initiated the evaluation of the model in a simulated phishing detection scenario. | Abhishek H, Aditya W | 50% |
| Real-world Deployment | Deployed the model in a controlled real-world environment and began monitoring its performance. | Muzammil S | 30% |
| Deployment Issue Resolution | Addressed minor deployment issues related to data ingestion and processing. | Muzammil S | 80% |

**Upcoming Tasks:**

| **Task Name** | **Description** | **Team Member Responsible** |
| --- | --- | --- |
| Deployment Monitoring | Continue monitoring the deployed model for phishing detection efficacy and system performance. | Muzammil S |
| User Feedback Collection | Initiate the collection of user feedback to identify areas for improvement in model performance and user interface. | Abhishek H |
| Model Fine-tuning | Fine-tune the deployed model based on feedback and performance evaluation. | Abhishek H |
| UI Enhancement | Enhance the user interface based on collected feedback to improve usability and functionality. | Muzammil S, Aditya W |

**Week 6-12: Monitoring, Feedback Collection, and Iteration**

**Tasks in Progress or Completed:**

| **Task Name** | **Description** | **Team Member Responsible** | **Percentage Complete** |
| --- | --- | --- | --- |
| Model Performance Monitoring | Daily monitoring of model performance to ensure stability and accuracy. | Muzammil S | Ongoing |
| User Feedback Collection | Collection of user feedback to identify areas for improvement in model performance and user interface. | Abhishek H | Ongoing |
| Model Iteration | Bi-weekly iterations on the model based on collected feedback and identified issues. | Muzammil S | Ongoing |
| UI Enhancement | Continuous enhancement of the user interface based on collected feedback to improve usability. | Muzammil S, Abhishek H | Ongoing |

**Upcoming Tasks:**

| **Task Name** | **Description** | **Team Member Responsible** |
| --- | --- | --- |
| Model Fine-tuning | Fine-tuning the model based on feedback and performance evaluation. | Aditya W |
| Data Augmentation | Implementing additional data augmentation techniques to improve model robustness. | Aditya W |
| Performance Evaluation | Conducting in-depth performance evaluation of the model to assess its efficacy and identify areas for improvement. | Muzammil S |
| Model Refinement | Refining the model architecture and parameters based on feedback and performance analysis. | Muzammil S, Aditya W |
| Documentation Update | Updating project documentation to reflect changes made during monitoring, feedback collection, and iteration. | Abhishek H |

**Final Weeks: Project Wrap-Up and Future Planning**

**Tasks in Progress or Completed:**

| **Task Name** | **Description** | **Team Member Responsible** | **Percentage Complete** |
| --- | --- | --- | --- |
| Project Presentation Preparation | Preparation for the final project presentation, including summarizing achievements, challenges, and key learnings. | Abhishek H | 100% |
| Drafting Final Project Report | Drafting the final project report detailing the methodology, model development, and deployment process. | Aditya W | 100% |
| Finalizing Project Presentation | Finalizing the project presentation, ensuring all key points are effectively communicated. | Muzammil S | 100% |
| Delivering Project Presentation | Delivering the final project presentation to stakeholders, demonstrating the project's success. | Aditya W | 100% |
| Project Retrospective Meeting | Conducting a project retrospective meeting to discuss achievements, areas for improvement, and future projects. | Project Lead | 100% |
| Documentation Update | Updating project documentation to include final project outcomes, learnings, and future recommendations. | Abhishek H | 100% |
| Post-Project Activities | Conducting post-project activities such as cleaning up project spaces, finalizing documentation storage, etc. | Aditya W, Abhishek H | 100% |

**Upcoming Tasks:**

| **Task Name** | **Description** | **Team Member Responsible** |
| --- | --- | --- |
| Future Planning Discussion | Holding discussions on potential future projects that could build on the current work and exploring new opportunities. | Muzammil S, Abhishek H , Aditya W |
| Follow-up Meeting Planning | Planning a follow-up meeting to revisit the project's impact after a few months and discuss any further developments. | Muzammil S, Abhishek H, Aditya W |