**Meeting Minutes 5**

**Meeting with Professor in class**

|  |  |
| --- | --- |
| **Item** | **Details** |
| Date and Time | 20-Mar-2025 |
| Location | In class (After the Presentation) |
| Attendees | Kanchan, Lakshmi, Sachin, Karthik, Narendran |
| Absent |  |
| Agenda | * Finalize the direction for the upcoming presentation. * Confirm that previously identified model enhancements (e.g., feature engineering, scaling) are on track. * Address any open concerns regarding data quality or runtime issues. * Outline tasks and responsibilities before next milestone |
| Discussion Points | * Quick updates on how each model (ARIMAX, SARIMAX, LSTM, Prophet, SVM, XGBoost) is performing. * Brief mention of any newly discovered literature or methods to improve model accuracy. * Agreement on how to present model performance comparisons. * Decision on highlighting successes and lessons learned. * Clarify any unresolved issues from earlier meetings (e.g., data pipeline or feature expansion). * Note the need for quality checks on the dataset if new sources have been added. * Timeline for completing final slides and rehearsal. * Confirmation of responsibilities (who updates which section of the presentation, who compiles the final deck, etc.). |
| Decisions Made | * A cohesive flow, starting with problem statement, dataset details, model exploration, and final outcomes. * Team agreed to emphasize both the improvements gained through feature engineering and scaling, as well as lessons learned about overfitting. |
| Actions Items | * **Kanchan**: Prepare initial draft slides for the presentation section on Time-Series analysis (LSTM). * **Lakshmi**: Finalize ARIMAX/SARIMAX performance metrics and gather results for comparison. * **Sachin**: Compile SVM results and coordinate the final deck layout. * **Karthik**: Update XGBoost performance metrics with any newly engineered features. * **Narendran**: Confirm Prophet results with reduced training period; provide summary slides on runtime reduction. |
| Any concerns (if any) | **Overfitting Check**: Ensure the updated models are fully tested (train vs. validation) to confirm overfitting is addressed.  **Time Constraint:** Maintain strict adherence to the timeline for final presentation prep. |
| Next Meeting Date and Time | 24-Mar-2025 |
| Minutes Prepared By | Naraendran Asokan |

|  |  |
| --- | --- |
| **Item** | **Details** |
| Date and Time | 24-Mar-2025 |
| Location | Virtually |
| Attendees | Kanchan, Lakshmi, Sachin, Karthik, Narendran |
| Absent |  |
| Agenda | * Discuss new insights from literature on model tuning and feature engineering. * Outline a plan for hyper-parameter tuning (e.g., grid search, random search, Bayesian optimization). * Assign responsibilities and set timelines for the next iteration of experiments. |
| Discussion Points | * Recent papers highlight the effectiveness of dropout regularization in time-series forecasting (particularly for LSTM). * Suggested combining rolling window cross-validation with hyper-parameter tuning for time-series models. * Additional metrics, such as Mean Absolute Percentage Error (MAPE), to better capture forecasting performance. * The team decided to test multiple search strategies (grid, random, and Bayesian) to find optimal parameters for each algorithm. * Early emphasis on LSTM, Prophet, and XGBoost, as they showed the greatest variability in performance. * A brief discussion on potential computational constraints; the team will consider selective subsets of data for certain time-consuming tuning processes. * Refine the dataset by applying newly discovered feature engineering techniques (e.g., seasonal decomposition for time-series, sentiment normalization). * Determine how to measure success: e.g., looking for at least a 5%–10% improvement in primary metrics before finalizing hyper-parameters. |
| Decisions Made | * Each team member will pick at least two recent publications to glean tuning strategies or advanced regularization methods. * A shared document will be created to summarize each paper’s key insights and how those might apply to our project. * The team agreed on a sequential approach: LSTM and XGBoost first, then Prophet, then SVM/ARIMAX/SARIMAX. * This sequence ensures the highest-impact models are optimized early, given finite time and computing resources. * Maintain a baseline set of metrics to gauge improvement consistently, including MAPE, RMSE, and R² for regression/time-series tasks. |
| Actions Items | * Kick off the initial hyper-parameter tuning for LSTM using grid search, log both training time and final metrics. * Investigate advanced dropout techniques or layer modifications from new literature. * Continue refining the ARIMAX/SARIMAX models using any new feature engineering insights from the literature survey. * Compile and share a summary of advanced statistical methods for time-series improvement. * Oversee the random search approach for SVM Regressor, including scaling variations. * Coordinate a shared Google Drive for all tuning logs and final results. * Lead the Bayesian optimization approach for XGBoost, focusing on tree-depth, learning rate, and regularization parameters. * Ensure version control for all experiment scripts in the team’s repository. * Perform runtime feasibility checks for Prophet’s tuning, especially on seasonality parameters. * Explore new ways to incorporate sentiment or external features into Prophet’s forecast. |
| Any concerns (if any) |  |
| Next Meeting Date and Time | 01-April-2025 |
| Minutes Prepared By | Narendran Asokan |

**Meeting with team members (Meeting with Ma’am)**

|  |  |
| --- | --- |
| **Item** | **Details** |
| Date and Time | 28-Mar-2025 |
| Location | Physical |
| Attendees | Kanchan, Lakshmi, Sachin, Karthik |
| Absent | Narendran Asokan |
| Agenda | * Status updates on model development and data preparation for stock predictions. * Review of company selection and forward plan. * Clarification about the decision regarding pipeline and frontend implementation. |
| Discussion Points | * Each team member provided a quick summary of current model-building efforts and dataset refinement. * Preliminary results for forecasting and sentiment analysis were shared. * Professor emphasized that the team should continue working with the Phase 1 companies: Meta, Apple, Snapchat, Google, and Microsoft. * These companies will form the core dataset for further experiments and analyses. * The professor confirmed that, for the time being, the team will not move forward with building a fully automated pipeline or a user-facing frontend. * Instead, the focus will remain on solidifying the analysis and improving model accuracy for the selected companies. * The team discussed potential improvements in hyper-parameter tuning for each ML model. * Feature engineering ideas (e.g., rolling averages) were briefly revisited to ensure thorough coverage across all five companies. |
| Decisions Made | * Meta, Apple, Snapchat, Google, and Microsoft remain the priority. * Any new methods or model improvements should be applied across these five datasets. * The professor directed the team to hold off on building an end-to-end pipeline or user interface. * Instead, the group’s immediate focus will be research-oriented: refining models, data preprocessing, and comparative evaluation. * The team will finalize model tuning strategies for each selected company, ensuring consistent methodology and metrics. |
| Actions Items | * Take ownership of one company’s dataset and run a thorough hyper-parameter tuning process for at least one time-series model (e.g., LSTM, Prophet). * Liaise with the overfitting lead to mitigate any persistent high variance issues. * Oversee another company’s dataset, focusing on ARIMAX/SARIMAX or SVM approaches. * Coordinate new strategies from the overfitting lead and apply them to reduce model variance. * Handle a third company; focus on advanced hyper-parameter optimization methods (random or Bayesian search). * Work with the overfitting lead to test more rigorous cross-validation techniques. * Manage a fourth company’s dataset with emphasis on XGBoost and feature engineering best practices. * Document tuning parameters and share results for group review. * Serve as the overfitting lead (tentative) and manage the fifth dataset. * Compile a short guide to overfitting remedies, including dropout strategies, scaling techniques, and cross-validation frameworks. |
| Any concerns (if any) |  |
| Next Meeting Date and Time | 01-April-2025 |
| Minutes Prepared By | Narendran Asokan |