Practicum Project - Spring 2025

Christian Robinson
crobinson321@gatech.edu

Project Sponsor: Cheyenne Herrera, SLTX Director of Operations

Abstract—This practicum project focuses on developing an interactive communication performance dashboard for the Surplus Lines Stamping Office of Texas (SLTX) Operations Department. While current internal metrics emphasize data entry and policy audits, communication is essential but underrepresented in performance metrics. This project uses data from Microsoft Outlook and Teams to quantify communication activity, categorized by method (email or phone), direction (outgoing or incoming), and content. A Shiny dashboard was developed to visualize communication patterns, highlight individual employee contributions, and provide management with actionable insights. To support deeper analysis, emails were also classified into six categories using natural language processing techniques. Three models—Multinomial Naive Bayes, Support Vector Machines (SVM), and Convolutional Neural Networks (CNN)— were compared using Monte Carlo cross-validation, with SVM achieving the highest accuracy. This dashboard not only enhances transparency in employee performance metrics, but also offers SLTX leadership new tools for data-informed decision-making.

1 INTRODUCTION

The Surplus Lines Stamping Office of Texas (SLTX) is a quasi-government organization responsible for oversight of the surplus lines insurance market in Texas. Per state law, each surplus lines licensed agent or broker producing business in Texas is required to report to SLTX data for every transaction it produces in the state. The Policy Analyst position in the Operations department is the backbone of the organization, tasked with three core responsibilities to assist agents in making timely filings and ensuring data quality, namely: data entry, policy audits, and communication with agents. Communication can involve assisting agents with compliance, support in navigating SLTX's filing application, providing end-of-period reports, or other miscellaneous items.

The project will focus on developing an interactive dashboard to quantify communication metrics for SLTX's Policy Analysts. While data entry and policy audits are quantified and used in performance evaluation, external communication is an underrepresented aspect of internal metrics yet is a vital aspect to Policy Analyst duties. The project will use Microsoft Outlook and Teams data to build a dashboard via a Shiny App in R. The data will undergo cleaning and preparation before being displayed in the application. Additionally, emails will be categorized based on their content using natural language processing models so that each can be routed and handled appropriately, so that management can gain further insight into the nature of ongoing communication and make decisions accordingly. Therefore, the main goals of the dashboard are to quantify employee performance and provide insight to management on communication patterns for future decision making.

2 PROBLEM STATEMENT

2.1 Dashboard Design

The primary deliverable for this project is a user-friendly Shiny dashboard that can integrate seamlessly with the existing internal performance dashboard used by SLTX Operations. The layout mirrors the current 2x2 grid (see Figure 1) and includes:

1. Communication counts separated by direction and method (phone, email)
2. A performance pie chart using the same visual logic as existing charts, and with interactive tooltips
3. A time-series comparison of volume by user with interactive tooltips
4. A bar charts showing email volume by user
5. A bar chart showing phone call count and duration by user
6. Within the bar chart showing email count by user, an interactive tooltip that summarizes email counts by email category, which is predicted using Natural Language Processing

The dashboard includes filtering options consistent with current internal tools and introduces interactivity to surface detailed insights. Specifically, the dashboard includes the ability to filter by a date range and by specific users.

2.2 Data Sources

Two primary data sources are used to support the dashboard:

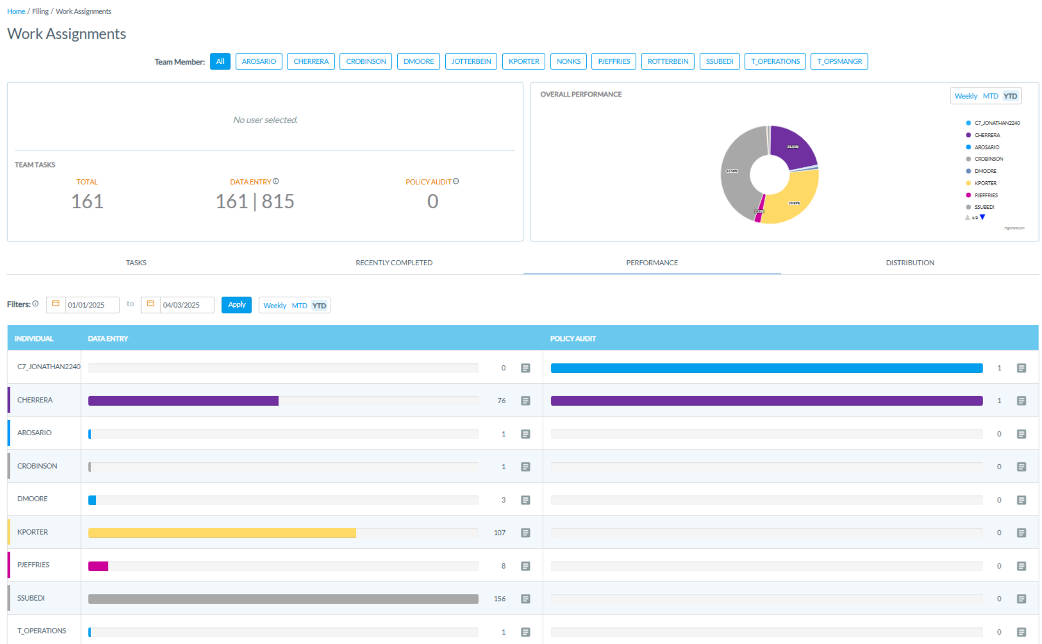

Figure 1—Current Work Assignments Dashboard

1. **Phone Data:** Microsoft Teams operator and user call logs. Logs only include individual call data for March 2024, July-November 2024, and January-February 2025. Additionally, operator logs from December 2024 were included. Current Microsoft and SLTX internal permission structures prevented the ability to obtain additional data for the purposes of this project. In all, phone data comprised of 247 calls. Logs were provided in separate csv files for each month.
2. **Email Data:** Extracted from shared Outlook folder using Python scripts to scrub relevant data points. Final extraction was performed on March 15, 2025 and included 15,731 readable emails over the preceding 12 months. The Body and Subject of each email was also extracted for use in downstream NLP tasks.

2.3 Data Preparation

For the dashboard, the phone and email data were cleaned and combined. The operator logs required joining 4 separate tables to obtain the ultimate end-point for each routed call and their corresponding call lengths. Then, the operator logs were combined with each csv for individual logs in R. For the email data, the following data points were extracted:

1. DateReceived- the date the email was received
2. Subject- the Subject line of the email

3. Body- the full Body text of the email
4. Category- the category assigned by SLTX staff for each email. Per internal processes, the email category is the user responsible for the email. Therefore, the extracted Category is used to obtain the user information for the dashboard
5. Subfolder- the location of the email within the shared inbox subfolder
6. Attachments- binary indicator for whether the email contains attachments
7. ChainID- a unique ID provided for each email chain to show emails of the same chain
8. Direction- binary indicator for the direction of the email; incoming or outgoing

Once extracted, the email data required substantial cleaning. Some emails contained embedded tables or other unusual characters that prohibited proper formatting of the body text. In these cases, the body content was split across numerous rows and columns rather than in a single field. Additionally, the text provided within embedded tables could not be reliably extracted. Emails that could not be extracted cleanly were excluded from the analysis, including training and evaluating the email classification models. Then, after cleaning the individual datasets, the phone and email data were combined into a unified dataset. Final preparation steps included filtering the data to retain only records associated with Operations staff in non-management roles. For the relevant staff, the colors that were used were obtained from SLTX's internal database and added to the final table.

In order to classify each email given its content, a category was first manually assigned to each email. In subsequent modeling steps, the model could then be trained accordingly and its accuracy reported against the manually provided category. For simplicity, there were 6 email categories assigned:

1. **Filing Assistance**- assigned to emails that pertained to *what* should be filed by the user. This includes assistance in interpreting policy documents, providing the policy data requiring reporting, or other items related to policy reporting.
2. **Technical Support**- assigned to emails that pertained to *how* to report. This includes guidance on navigating SLTX's filing application, debugging errors, or assisting users in login issues.
3. **Audits**- assigned to emails that pertained to audits performed by SLTX. SLTX performs ongoing audits of users to ensure data quality and compliance, and Audit emails are those pertaining to these requests, including providing requested documents and ensuring correction of noncompliant items discovered in the audit process.

4. **Reports**- assigned to emails relating to SLTX reports, both those provided individually to users and those provided to other interested stakeholders, such as the Texas Department of Insurance.
5. **API Mapping**- assigned to emails pertaining to assistance provided to users who report data to SLTX via API.
6. **Other**- assigned to emails that do not fit in any of the categories listed above.

In all, the final dataset included 11 data points and 10,947 items (10,631 emails; 316 calls) across 9 users from February 1, 2024 - March 15, 2025.

3 METHODOLOGY

3.1 Dashboard

The dashboard was developed in R using the Shiny framework to display communication metrics from Microsoft Teams and Outlook. It mirrors the layout and style of SLTX's existing internal dashboards and includes the following components:

1. **Summary table** displaying total communications across selected filters, broken down by:
   (a) Total Communications
   (b) Total Phone Calls (Incoming/Outgoing)
   (c) Total Emails (Incoming/Outgoing)
   Each row includes contextual tooltips and coachmarks to aid interpretation.
2. A **line chart** displaying total communications over time by user. If the dashboard user selects a time frame greater than 1 month, then the chart groups totals by month. Otherwise, the chart shows the daily totals. The line chart also includes a tooltip for each data point which, when hovered over, displays the time period (Month/Year or Date), the user name, the phone call total, email total, and overall total.
3. A **donut chart** visualizing the percentage of total communications by user. Hovering displays each user's name, total counts, and percent contribution.
4. A **bar chart** ranks users by total email volume. Each bar includes:
   (a) Two interactive icons:
      i. Green Triangle: filters to outgoing emails
      ii. Red Triangle: filters to incoming emails
   (b) When clicked, a table appears summarizing predicted email categories for that user. A "Show More" option expands into an interactive dataframe pertaining to user emails.

Category	Count
Technical Support	80
Reports	76
Filing Assistance	56
Audits	2
Other	2
TOTAL	216

Figure 2—Email Count by Category for Single User

5. A **bar chart** showing phone call count by user. Includes:
   (a) Toggle to switch between call count and total duration
   (b) Tooltip per bar showing user name, total calls, total duration, and average call length

The filters are defaulted to all users for the date range January 1, 2025 to present.

3.2 Email Classification Model

To create the email classification model which assigns a category for each email, the following models were tested to compare accuracy and overfitting: **Multinomial Naive Bayes, SVM, and CNN**. Multinomial Naive Bayes was chosen as a base model due to its relative simplicity. SVM was chosen because, while also relatively simple, it is also able to capture more complex relationships than Naive Bayes. Finally, CNN was chosen as the neural network model of choice. The encoding method chosen for each is dependent on the model being used; for Naive Bayes and SVM, TF-IDF was used, for CNN, GloVe was used for embedding.

Additionally, models were created for predicting *Category* given the email *Subject* and the *Body*. Before modeling, the dataset was broken up into a training set and test set using a 70/30 split. To find the optimal model, the following process was applied:

1. Hyperparameter tuning on the training data set using cross-validation in a grid search.
2. Created final model using optimal hyperparameters for each model trained
3. Deployed final models on the test set. 40-fold Monte Carlo Cross Validation was used to report accuracy metrics.

3.2.1 *Exploratory Data Analysis*

To understand the dataset prior to modeling, we first examined category distribution. Filing Assistance, Reports, and Technical Support dominate the dataset, necessitating stratified sampling when splitting the data.

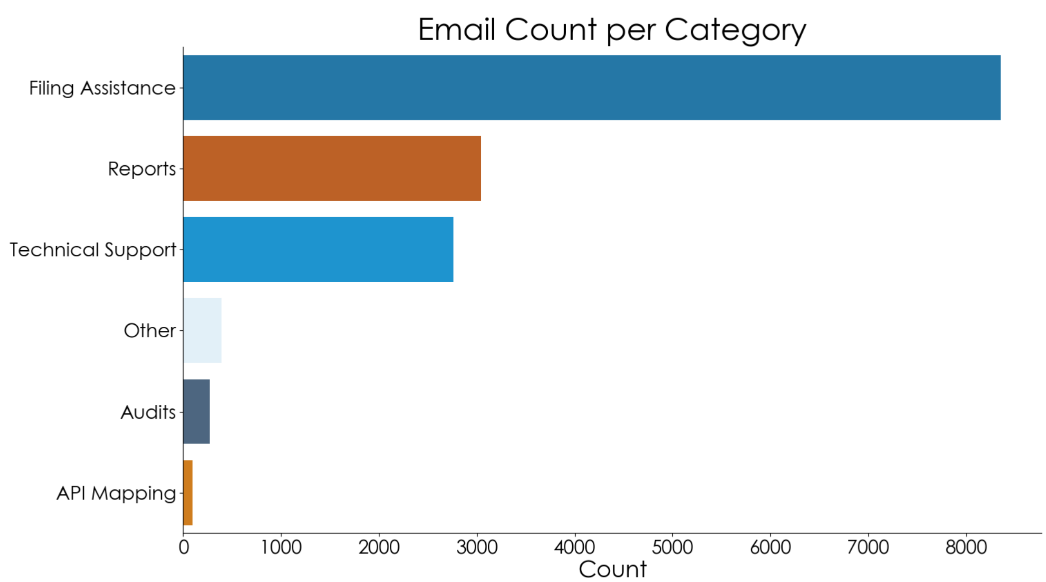

Figure 3—Email Count by Manually Assigned Category

Principal Component Analysis (PCA) was used to visualize separability between categories. As shown in Figure 4, Reports and API Mapping stand apart, while other categories cluster around the origin—indicating potential difficulty in distinguishing them. The first two components account for only 11.2% of total variance.

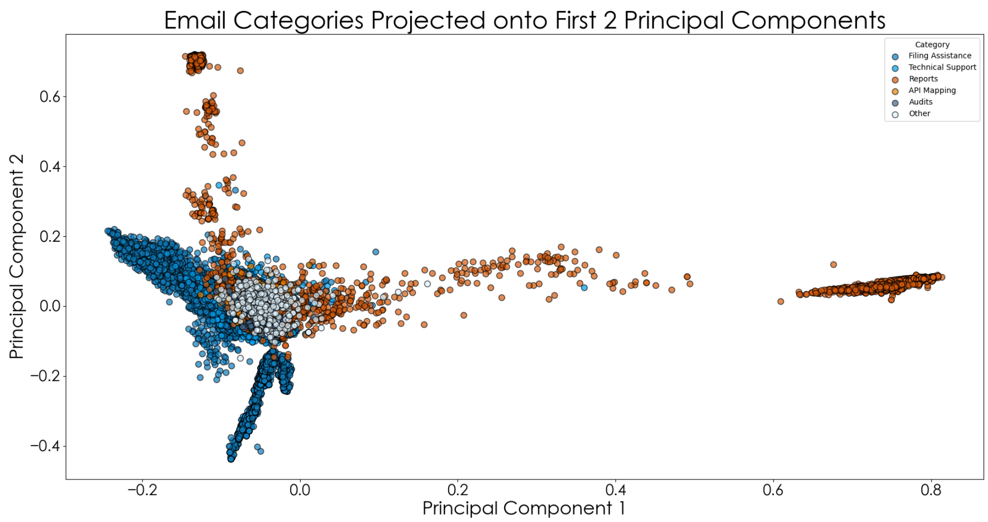

Figure 4—Scatter Plot of First 2 Principal Components by Category

We also examined the structure and content of emails. Boxplots (Figures 5 and 6) show that API Mapping emails tend to have consistent body lengths, while Technical Support emails often feature longer subject lines. Most categories exhibit a right-skewed distribution in subject line length.

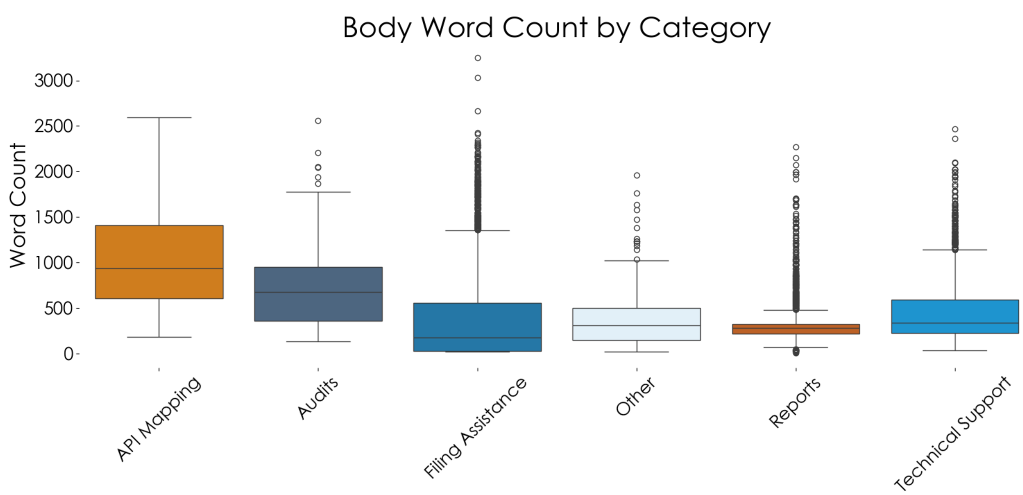

Figure 5—Boxplot of Word Count in Email Body for Training Set

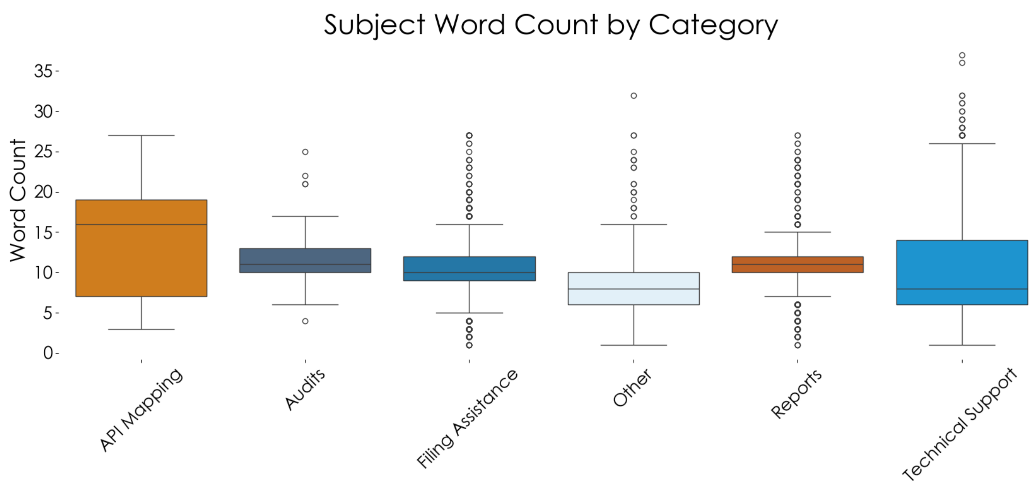

Figure 6—Boxplot of Word Count in Subject Line for Training Set

Regarding content, word frequency analyses (Figures 7 and 8) reveal some terms are broadly used (e.g., "insurance", "sltx", "policy"), while others are more category-specific (e.g., "audit", "report", "api"). Interestingly, subject lines provide more discriminative terms than body text. For example, "policy" is strongly associated with Audit emails when in the subject line, but is common across categories when appearing in the body.

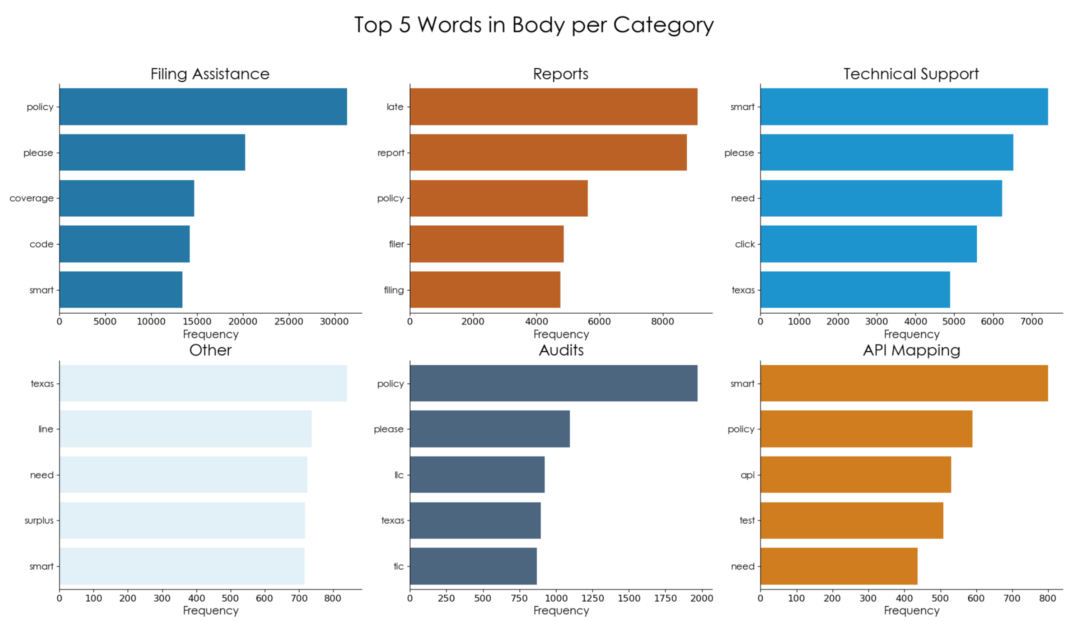

Figure 8—Top 5 Words in Body per Category for Training Set

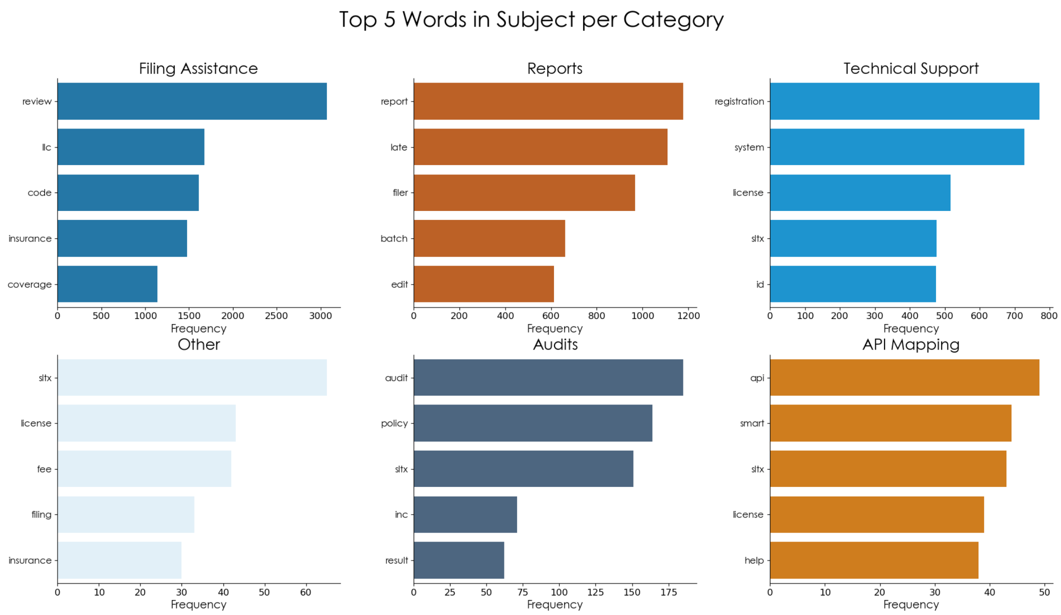

Figure 7—Top 5 Words in Subject Line per Category for Training Set

3.2.2 *Hyperparameter Tuning*

In the first modeling step, the training set is used to find the optimal hyperparameters which balance accuracy on the training set with overfitting. To determine the best hyperparameters, a grid search was employed with 10-fold cross validation, where the accuracy was recorded for each fold and set of hyperparameters. Then, considering the mean and standard deviation of the accuracy for each model, optimal sets of hyperparameters were chosen to be used in the final models. The mean and standard deviation are both considered to prioritize a balance of accuracy and consistency.

Beginning with the Multinomial Naive Bayes model, the only hyperparameter tuned is **alpha**, which is a smoothing parameter. As shown in Figure 9 below, the optimal alpha is $\alpha = 0.1$, where the corresponding average accuracy across all folds was 91.2%. Additionally, the standard deviation of accuracy for each α was fairly consistent, ranging from 0.006 for $\alpha = 0.01$ to 0.008 for $\alpha = 0.3$. Therefore, the optimal α for the Multinomial Bayes Model is 0.1.

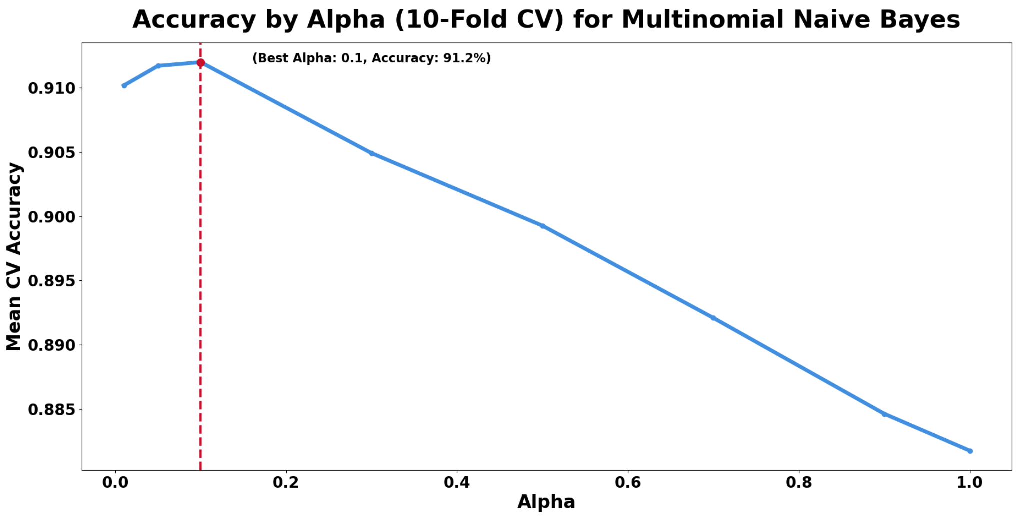

Figure 9—Hyperparameter Tuning Results on Training Set for Multinomial Naive Bayes

Next, optimizing the SVM model with a linear kernel requires tuning the hyperparameter $C = \mathtt{Cost}$. The hyperparameter controls the width of the margin and, therefore, represents the trade-off between accuracy and overfitting. As shown in Figure 10, the optimal hyperparameter in terms of accuracy is $C = 10$ where the corresponding accuracy was 95.2%. However, $C = 1$ had an accuracy of 95.1% and, while slightly lower than the optimal hyperparameter when evaluated on the training set, it better accounts for overfitting. Therefore, $C = 1$ is the proper hyperparameter to choose for our final SVM model.

To optimize the Convolutional Neural Network (CNN) model, the following hyperparameters were tuned:

1. Learning Rate: 0.001, 0.005
2. Max Epochs: 5, 10, 20
3. Batch Size: 32, 64
4. Activation Function: ReLU, Tanh, ELU
5. Number of Convolutional Layers: 1, 2
6. Window Sizes: (2,3), (3,4,5)
7. Hidden Dimensions: 50, 100, 150

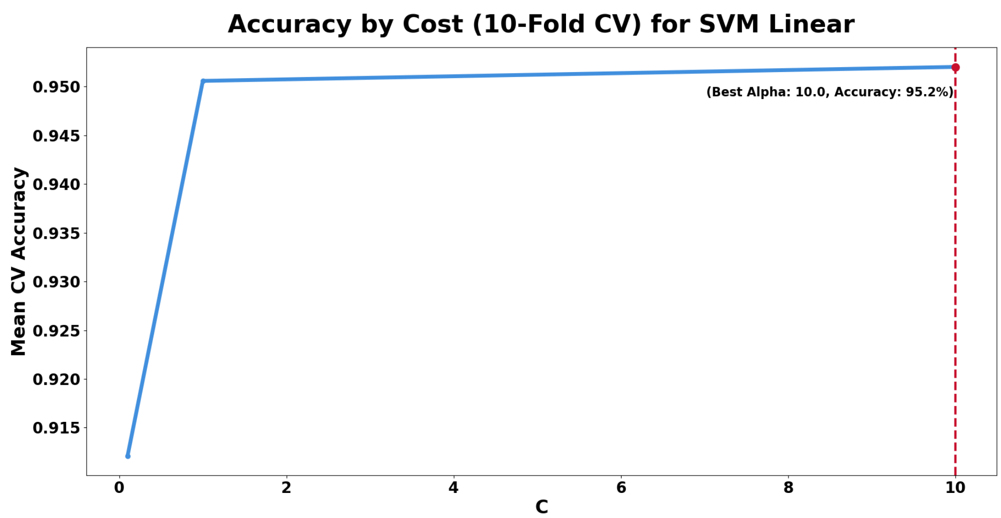

Figure 10—Hyperparameter Tuning Results on Training Set for SVM

After tuning the CNN model for each of the hyperparameters listed above, the optimal combination chosen when considering mean and standard deviation of accuracy is: **Learning Rate = 0.005, Max Epochs = 10, Batch Size = 32, Activation Function = ELU, Window Size = (3,4,5), and Hidden Dimensions = 50**. Figure 11 highlights how, though the mean accuracy is greater for max epochs = 20 for many combinations of hyperparameters, it also tended to overfit, with max epochs = 10 providing a better balance of accuracy while accounting for overfitting. Additionally, with the optimal activation function being ELU, this indicates that the feature set is likely either sparse or has high variance which enables the ELU activation function to better capture subtle patterns with a smoother gradient. Lastly, the optimal learning rate suggests the model benefits from faster updates during training.

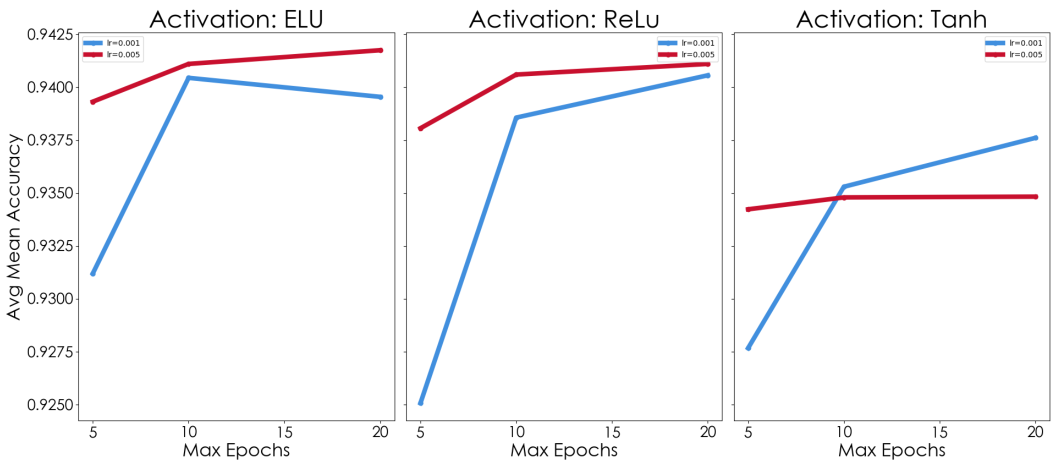

Figure 11—CNN Hyperparameter Tuning

4 ANALYSIS AND RESULTS

4.1 Dashboard

The final dashboard successfully integrates all of the required goals: metricizing communication production for Operations Department staff in a flexible, user-friendly dashboard. The left-hand panel places all filters and the legend in a central location to make interaction seamless. The use of tooltips and embedded tables fosters an environment where there is vast information without overwhelming the user, providing a clean look (Figure 12). The appendix shows detailed screenshots of each of the final charts in the dashboard.

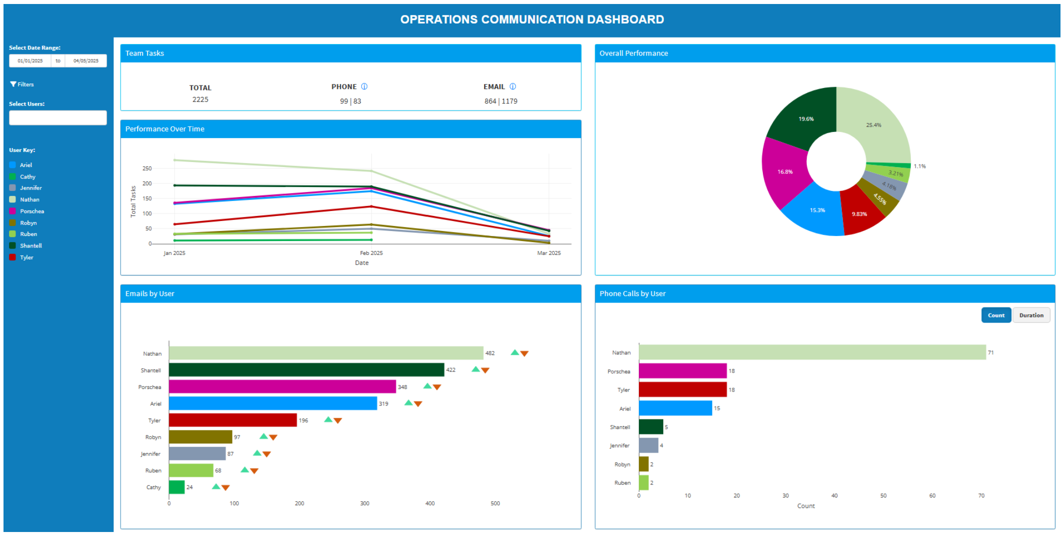

Figure 12—Final Dashboard

4.2 Email Classification Model

Monte Carlo Cross-Validation (40 iterations) was used to evaluate model stability and generalization. The SVM model with a linear kernel consistently outperformed both the Multinomial Naive Bayes and CNN models in terms of average accuracy and variance (Table 1):

Table 1—Monte Carlo Cross Validation: Testing Accuracy by Model

Statistic	Naive Bayes	SVM Linear	CNN
Mean	0.913726	0.953272	0.944611
Variance	0.000012	0.000007	0.0.000020

SVM's high accuracy and low variance make it a reliable candidate for production use. Moreover, the model's speed and interpretability offer practical advantages over the CNN.

5 CONCLUSIONS

This project successfully delivered a flexible, user-friendly Shiny dashboard that quantifies and visualizes external communication for SLTX Policy Analysts—filling a critical gap in internal performance evaluation tools. By integrating Microsoft Teams and Outlook data, the dashboard empowers both individual staff and leadership with new visibility into operational workload and responsiveness. Importantly, the addition of natural language processing to classify emails adds context to communication metrics and offers a richer understanding of how staff time is allocated across filing assistance, audits, technical support, and other key categories.

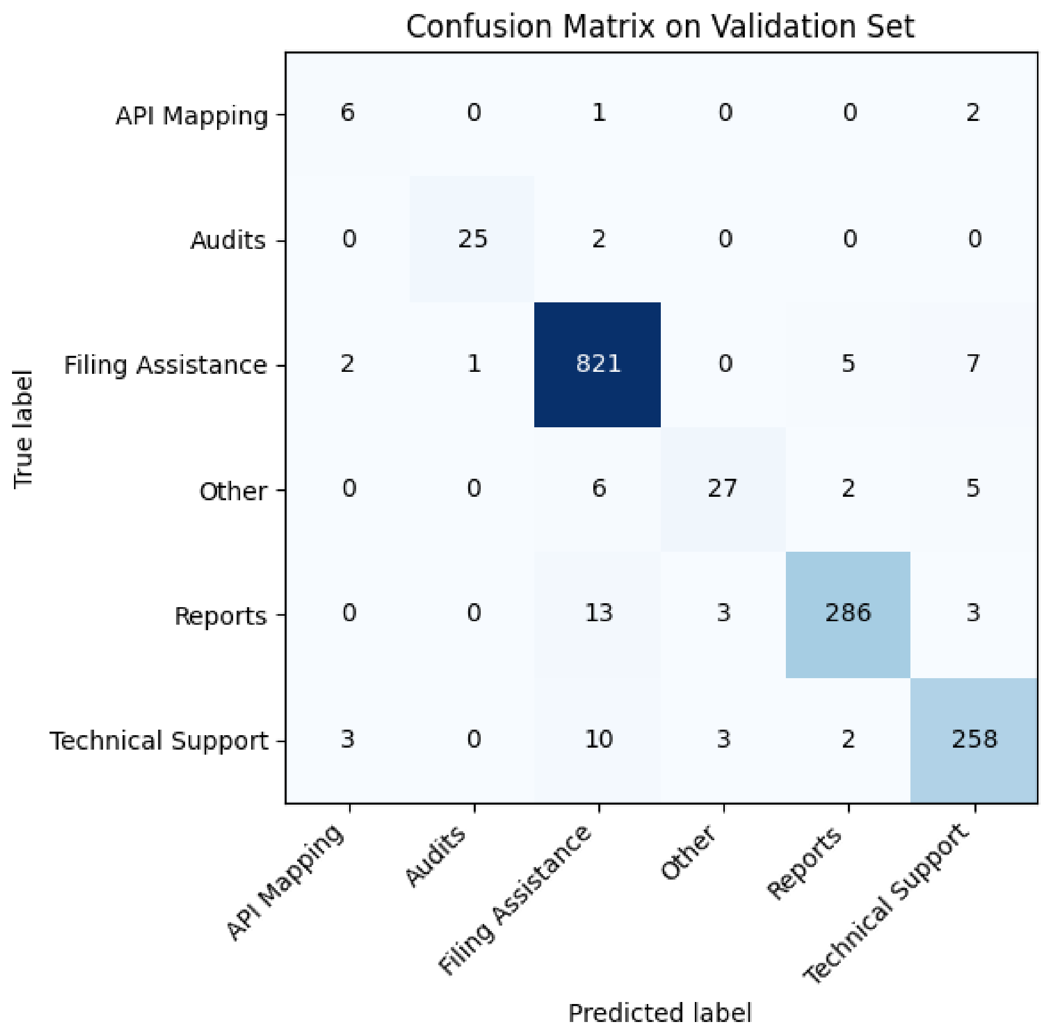

Figure 13—Confusion Matrix for Predictions Made with Final SVM Model

The email classification model evaluation revealed that the SVM with linear kernel provided the most accurate and stable results, outperforming both the simpler Multinomial Naive Bayes and the more complex CNN. This model can be used to support real-time email categorization in future iterations. To further validate its generalizability, the model was evaluated on a separate holdout validation set containing 1,493 emails. The confusion matrix in Figure 13 shows the results of the final model on the validation set.

As the confusion matrix shows, the categories with fewer observations- API Mapping, Audits, and Other- have the lowest accuracy, though each still performed well. Additionally, Filing Assistance has the most false positives, which is likely due to it also having the most observations.

In the end, the limited availability of complete phone call records remains a notable constraint. For full functionality, SLTX should consider working with Microsoft to obtain continuous access to Teams call logs or develop an API-based solution for real-time data ingestion. Additionally, transitioning from static CSV-based reports to live data feeds would allow the dashboard to evolve from a retrospective tool into a proactive decision-support system. Continued collaborations will be essential for scaling this tool beyond its current prototype and maximizing its organizational value.

Appendix A: Final Dashboard Images

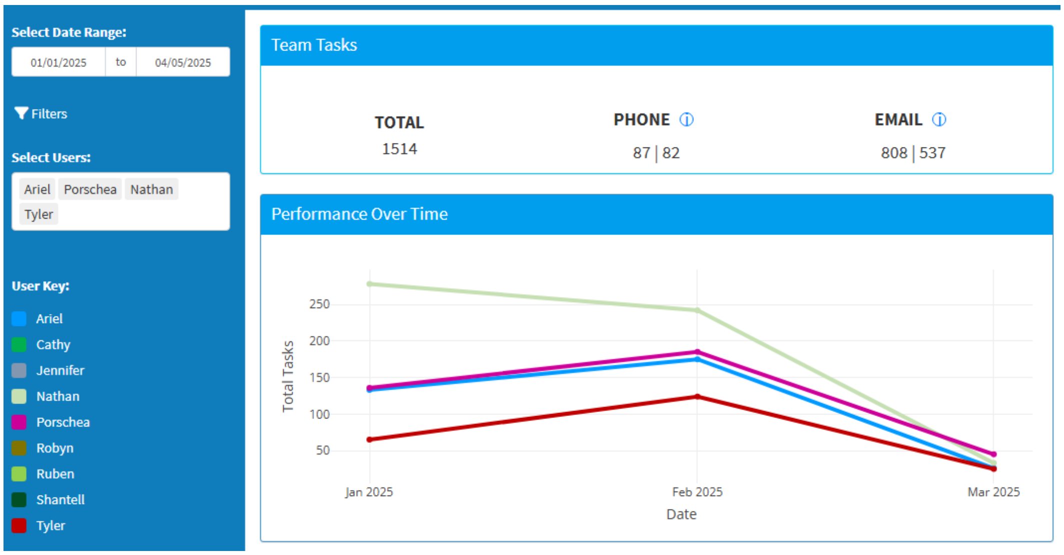

Figure 14—Final Dashboard- Total Tasks and Line Chart (View for Greater than 1 Month)

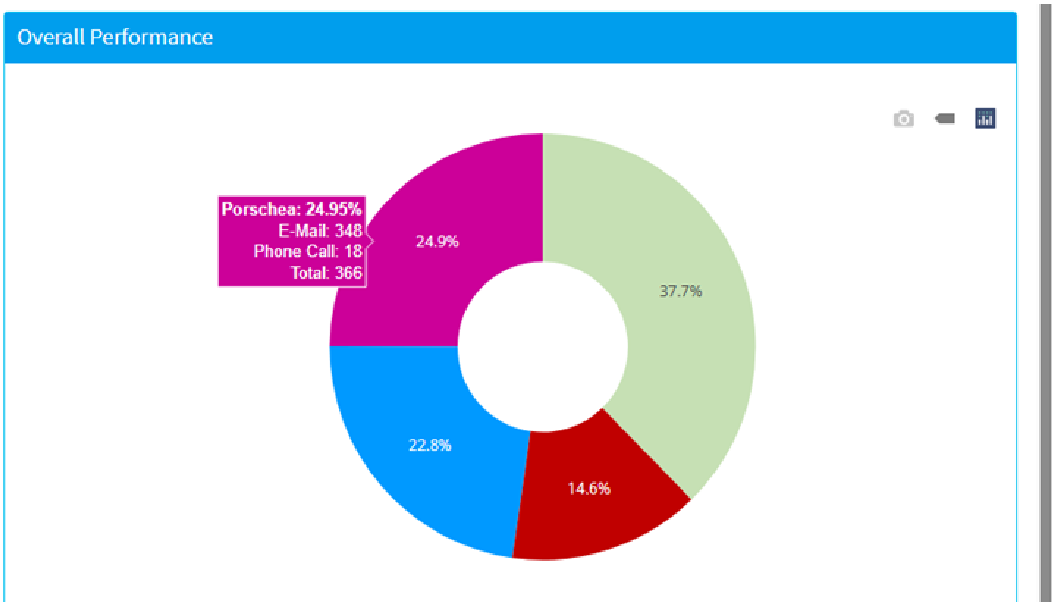

Figure 15—Final Dashboard- Donut Chart

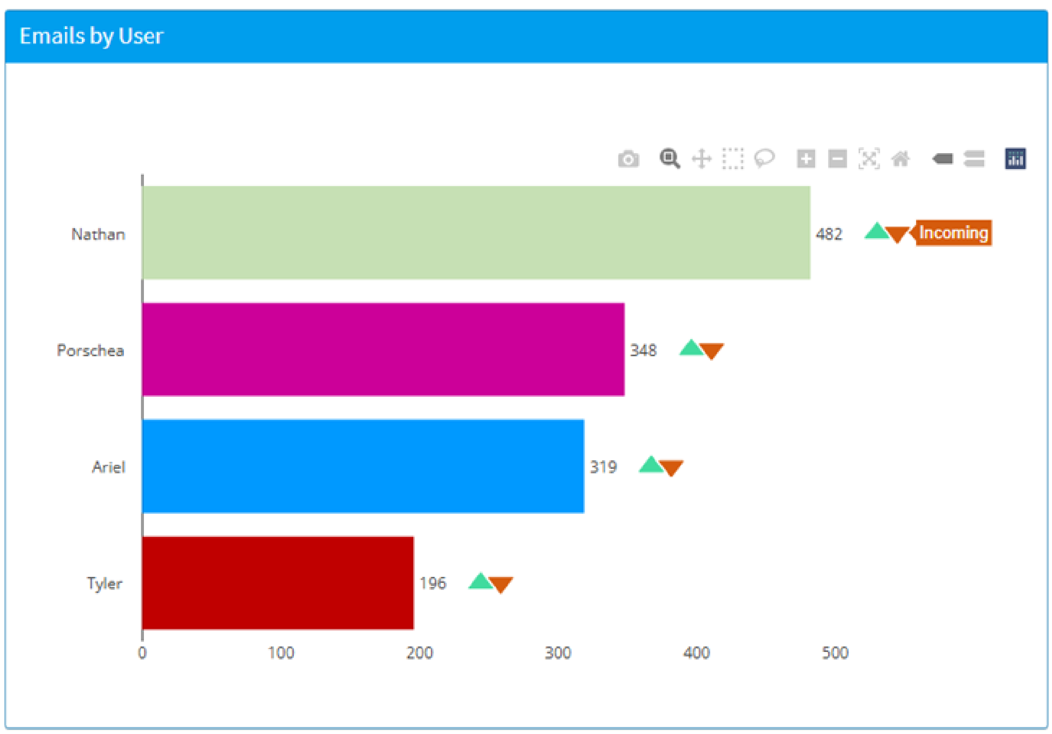

Figure 16—Final Dashboard- Email Bar Chart

Detailed Emails for Nathan

Show 10 entries

Search:

	Date	Subject	Body	Attachments	Category
1	03/09/2025	1894 - PCF Insurance Services of the West, LLC - 2452319	Hello, I am needing the coverage code for the attached [redacted]	Y	Filing Assistance
2	03/09/2025	COSTANZA INSURANCE AGENCY INC (15386) - Review Pol# [redacted]	Agent: COSTANZA INSURANCE AGENCY INC - 15386 User: KCARTER - KRYSTLE CARTER KRYSTLE CARTER has submitted a Renewal Policy [redacted]	N	Filing Assistance
3	03/07/2025	LANCE ALAN LAGERE (2244203) - Review Pol# [redacted]	Agent: LANCE ALAN LAGERE - 2244203 User: GKASHWER - GRAYSON KASHWER GRAYSON KASHWER has submitted a New Policy for policy [redacted]	N	Filing Assistance

Figure 17—Final Dashboard- Email Data Frame After Selecting "Show More"

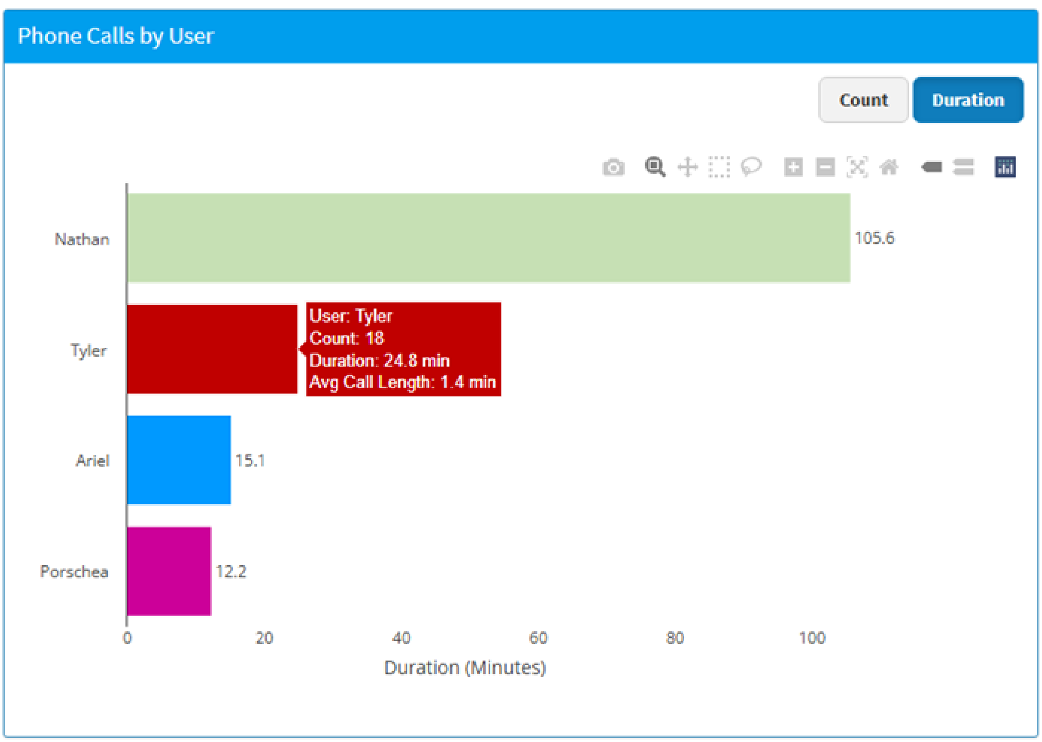

Figure 18—Final Dashboard - Phone Bar Chart

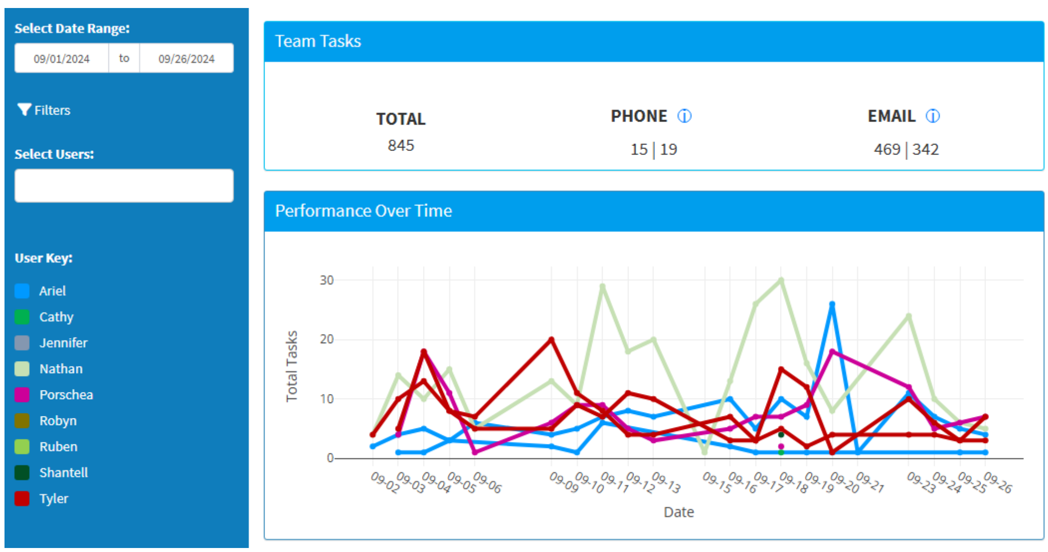

Figure 19—Final Dashboard- Line Chart (View for Less than 1 Month)