

Helpdesk Ticketing Support System Based on Natural Language Processing for a Government Organisation

Author: Abigail Davies
Supervisor: Dr. Seena Joseph

Project submitted as part of the requirements for the award of
BSc (Hons) Computing (Software Engineering)



Prifysgol Cymru
Y Drindod Dewi Sant
University of Wales
Trinity Saint David

1. Context of Project

Currently many organisations use helpdesk ticketing support systems to raise, track and deal with issues. The process allows an individual to raise a ticket which is sent to a department that analyses and redirects to the relevant department to deal with. This study investigates how natural language processing (NLP) can be utilised to automate the process. This will reduce human error, improving customer satisfaction and free employee time to action other important tasks.

2. Aim and Objectives

To develop a helpdesk ticketing support system based on natural language processing for a public organisation.

1. To critically analyse the current literature to understand the use of natural language processing in customer service and helpdesk systems.
2. To design and develop a model based on natural language processing for a helpdesk system to automate the redirection of tickets to the relevant departments.
3. To evaluate and compare the performance of the developed model using well-known evaluation metrics such as F1 score, accuracy, precision and recall.
4. To develop a web application that integrates the evaluated model for human interaction.

3. Literature Review Findings

The key findings:

- Studies of a similar size, use a dataset of approximately 2000 tickets on average.
- Balanced dataset = Greater results
- Various NLP techniques should be implemented in the pre-processing of data stage to allow the model to interpret data correctly.
- Support vector machines (SVM – figure 1) is the most popular traditional learning method for classification tasks.

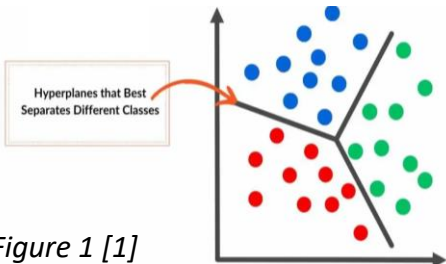
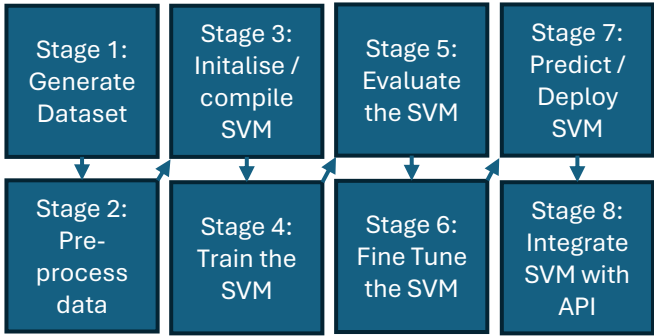


Figure 1 [1]

4. Methodology



5. Evaluation

The SVM was evaluated using several metrics, e.g. accuracy, precision, recall and F1-Score. Together they provide a clear view of the model's performance, comparing the actual results with the predicted results. The SVM scored 99.75% across all. The confusion matrix (figure 2) is a visual representation of the predictions.

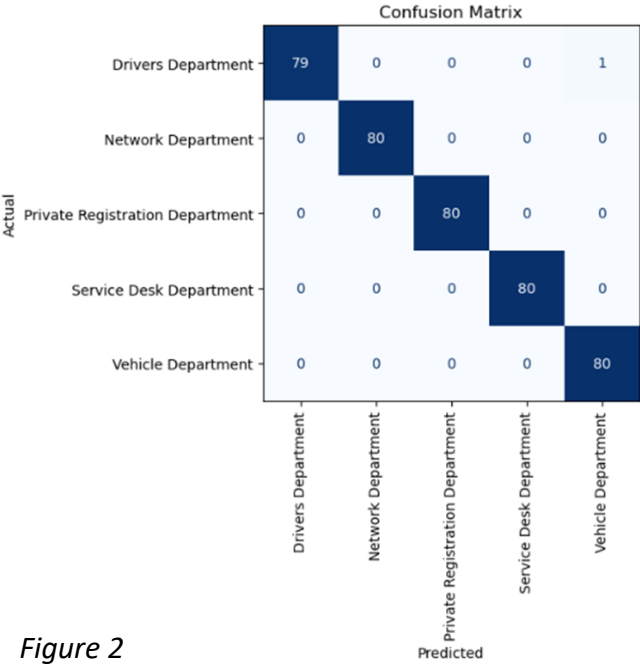


Figure 2

6. Conclusion

2000 tickets were generated for the dataset; 1600 tickets trained the SVM, and 400 tickets tested the SVM. 399 out of 400 tickets were predicted correctly. The model was successfully integrated with the API where new data was submitted, and the correct prediction was returned to the user.

7. Recommendations

Investigate model's performance if:

- Larger dataset was used.
- Real dataset was used.
- Dataset contained poorly constructed data.

8. References

[1] N.V. Otten, *SpotIntelligence*. May 6, 2024. [Online]. Available: <https://spotintelligence.com/2024/05/06/support-vector-machines-svm/>