Automatic Ticket Assignment

Capstone Project

Interim Report

AIML

# Index

[Index 2](#_Toc93758869)

[Team Details 3](#_Toc93758870)

[Summary of the problem statement, Data and findings 4](#_Toc93758871)

[Problem Statement 4](#_Toc93758872)

[Abstract 4](#_Toc93758873)

[Data & Findings 4](#_Toc93758874)

[Data provided in format 4](#_Toc93758875)

[Total Records 4](#_Toc93758876)

[Data Fields 4](#_Toc93758877)

[Sample data 4](#_Toc93758878)

[Distribution of classes 5](#_Toc93758879)

[Summary of the approach to EDA and Pre-Processing 6](#_Toc93758880)

[Cleaning processes applied 6](#_Toc93758881)

[Visualization 6](#_Toc93758882)

[Word Association 6](#_Toc93758883)

[Collapsible tree 6](#_Toc93758884)

[Word Cloud of entire data and top 3 categories 6](#_Toc93758885)

[Entire Data 6](#_Toc93758886)

[Charts 6](#_Toc93758887)

[Bokeh 6](#_Toc93758888)

[Barchart 6](#_Toc93758889)

[Decide Model and Model building 7](#_Toc93758890)

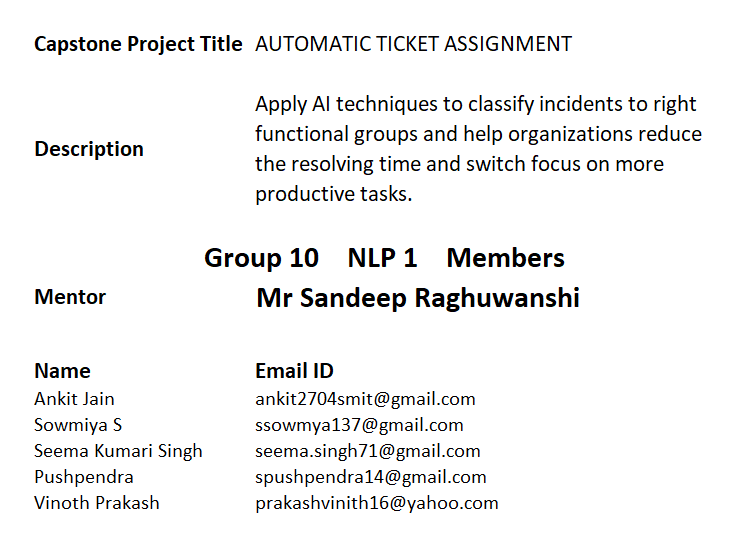
[Model performance - Approaches to improve model 8](#_Toc93758891)

[Code Snippet 9](#_Toc93758892)

[Finalized results 10](#_Toc93758893)

[Link to code and references 11](#_Toc93758894)

# Team Details



# Summary of the problem statement, Data and findings

## Problem Statement

IT leverages Incident Management process to ensure there is no disruption to business operations. Any unplanned disruption can cause interruption to business services. Incident management process helps in identification of issues or problems faced by users or operation teams. Manual assignment of incidents is time consuming and requires human intervention. There may be lag in incident resolution due to human errors or if incidents are not routed appropriately. On the other hand, manual assignment also increases the response and resolution times which result in user satisfaction deterioration / poor customer service.

## Abstract

Applying traditional machine learning and neural network-based NLP to automatically classify tickets and directing them to appropriate groups at the earliest, helps in improving the throughput in the ticketing pipeline of an organization.

## Data & Findings Data format CSV

### Total Records 8500

### Data Fields

|  |  |
| --- | --- |
| Short description | A brief overview of the issue faced by the user |
| Description | Detailed description of the issue |
| Assignment group | GRP\_0 ~ GRP\_73 (total 74 classes of Assignment group) |

### Sample data

| **Short description** | **Description** | **Assignment group** |
| --- | --- | --- |
| login issue | -verified user details.(employee# & manager name)  -checked the user name in ad and reset the password.  -advised the user to login and check.  -caller confirmed that he was able to login.  -issue resolved. | GRP\_0 |
| Outlook | received from: hmjdrvpb.komuaywn@gmail.com  hello team,  my meetings/skype meetings etc are not appearing in my outlook calendar, can somebody please advise how to correct this?  kind | GRP\_0 |
| cant log in to vpn | received from: eylqgodm.ybqkwiam@gmail.com  hi  i cannot log on to vpn  best | GRP\_0 |

### Distribution of classes

1. High imbalance seen in data with GRP\_0 having 40%+ percent of representation
2. Many groups/classes are with very little representation.
3. Null values:

Short description 8

Description 1

Assignment group 0

1. Certain values for the “Short descriptions” are same as “Description”
2. Very few ticket have Non-English descriptions

# Summary of the approach to EDA and Pre-Processing

### Cleaning processes applied

1. Removal of trailing spaces
2. Removal of line breaks and tabs (\r\n\t)
3. Removal of special characters
4. Removal of extra spaces
5. Missing value imputation: NLP keyword extraction with Rake < TODO >

## Visualization

### Collapsible tree

## Word Cloud of entire data and top 3 categories

### Entire Data

TBD

## Charts

TBD

# Decide Model and Model building

TBD

# Model performance - Approaches to improve model

TBD

# Code Snippet

# Finalized results

# Link to code and references