greatlearning

AIML

CAPSTONE PROJECT

NATURAL LANGUAGE PROCESSING

AUTOMATIC TICKET ASSIGNMENT

PROBLEM STATEMENT

- **DOMAIN:** Automatic Ticketing System.

- **CONTEXT:** One of the key activities of any IT function is to "Keep the lights on" to ensure there is no impact to the Business operations. IT leverages Incident Management process to achieve the above Objective. An incident is something that is unplanned interruption to an IT service or reduction in the quality of an IT service that affects the Users and the Business. The main goal of Incident Management process is to provide a quick fix / workarounds or solutions that resolves the interruption and restores the service to its full capacity to ensure no business impact. In most of the organisations, incidents are created by various Business and IT Users, End Users/ Vendors if they have access to ticketing systems, and from the integrated monitoring systems and tools. Assigning the incidents to the appropriate person or unit in the support team has critical importance to provide improved user satisfaction while ensuring better allocation of support resources. The assignment of incidents to appropriate IT groups is still a manual process in many of the IT organisations. Manual assignment of incidents is time consuming and requires human efforts. There may be mistakes due to human errors and resource consumption is carried out ineffectively because of the misaddressing. On the other hand, manual assignment increases the response and resolution times which result in user satisfaction deterioration / poor customer service.

- **Business Domain Value:** In the support process, incoming incidents are analysed and assessed by organisations support teams to fulfill the request. In many organisations, better allocation and effective usage of the valuable support resources will directly result in substantial cost savings. Currently the incidents are created by various stakeholders (Business Users, IT Users and Monitoring Tools) within IT Service Management Tool and are assigned to Service Desk teams (L1 / L2 teams). This team will review the incidents for right ticket categorization, priorities and then carry out initial diagnosis to see if they can resolve. Around ~54% of the incidents are resolved by L1 / L2 teams. Incase L1 / L2 is unable to resolve, they will then escalate / assign the tickets to Functional teams from Applications and Infrastructure (L3 teams). Some portions of incidents are directly assigned to L3 teams by either Monitoring tools or Callers / Requestors. L3 teams will carry out detailed diagnosis and resolve the incidents. Around ~56% of incidents are resolved by Functional / L3 teams. Incase if vendor support is needed, they will reach out for their support towards incident closure. L1 / L2 needs to spend time reviewing Standard Operating Procedures (SOPs) before assigning to Functional teams (Minimum ~25-30% of incidents needs to be reviewed for SOPs before ticket assignment). 15 min is being spent for SOP review for each incident. Minimum of ~1 FTE effort needed only for incident assignment to L3 teams. During the process of incident assignments by L1 / L2 teams to functional groups, there were multiple instances of incidents getting assigned to wrong functional groups. Around ~25% of Incidents are wrongly assigned to functional teams. Additional effort needed for Functional teams to re-assign to right functional groups. During this process, some of the incidents are in queue and not addressed timely resulting in poor customer service. Guided by powerful AI techniques that can classify incidents to right functional groups can help organizations to reduce the resolving time of the issue and can focus on more productive tasks.

- **PROJECT DESCRIPTION:**
  In this capstone project, the goal is to build a classifier that can classify the tickets by analyzing text.

  Pre-Processing, Data Visualisation and EDA
  - Exploring the given Data files
  - Understanding the structure of data
  - Missing points in data
  - Finding inconsistencies in the data
  - Visualising different patterns
  - Visualising different text features
  - Dealing with missing values
  - Text preprocessing
  - Creating word vocabulary from the corpus of report text data
  - Creating tokens as required

  Model Building
  - Building a model architecture which can classify.
  - Trying different model architectures by researching state of the art for similar tasks.
  - Train the model.
  - To deal with large training time, save the weights so that you can use them when training the model for the second time without starting from scratch.

  Test the Model, Fine-tuning and Repeat
  - Test the model and report as per evaluation metrics
  - Try different models
  - Try different evaluation metrics
  - Set different hyper parameters, by trying different optimisers, loss functions, epochs, learning rate, batch size, checkpointing, early stopping etc..for these models to fine-tune them
  - Report evaluation metrics for these models along with your observation on how changing different hyper parameters leads to change in the final evaluation metric.

- **PROJECT OBJECTIVE:**

  - Learn how to use different classification models.
  - Use transfer learning to use pre-built models.
  - Learn to set the optimisers, loss functions, epochs, learning rate, batch size, checkpointing, early stopping etc
  - Read different research papers of given domain to obtain the knowledge of advanced models for the given problem.

- **PROJECT SUBMISSIONS AND EVALUATION CRITERIA:**
  While we encourage peer collaboration and contribution, plagiarism, copying the code from other sources or peers will defeat the purpose of coming to this program. We expect the highest order of ethical behaviour.

  Submit the project as given below.
  - Report with Problem Statement, Related Work, Your Approach and comparison of results with other models written in PDF.
  - Github link where the model is hosted.

POINTS TO REMEMBER

1. A maximum of 60 points will be awarded for this project

2. Project to be submitted within 6 weeks of date of release. Late submission will be accepted under genuine situation. Score will be given as per the below formula:

   If the current score is greater than 40 then the final score will be capped at 40.

   Else the current score will be awarded.

3. Any form of plagiarism is strictly prohibited. No score will be awarded in this case.

HAPPY LEARNING!