

DARPG-Hackathon

Data Driven Innovation For Citizens Grievance Redressal

Ankita Bhatia
Siddharth Singh

Category

- **Category:**

Making the redressal process more robust and data-driven to reduce the Grievance submission and resolution lifecycle. Technology such as AI and ML could be used.

Velocity of complaint resolution

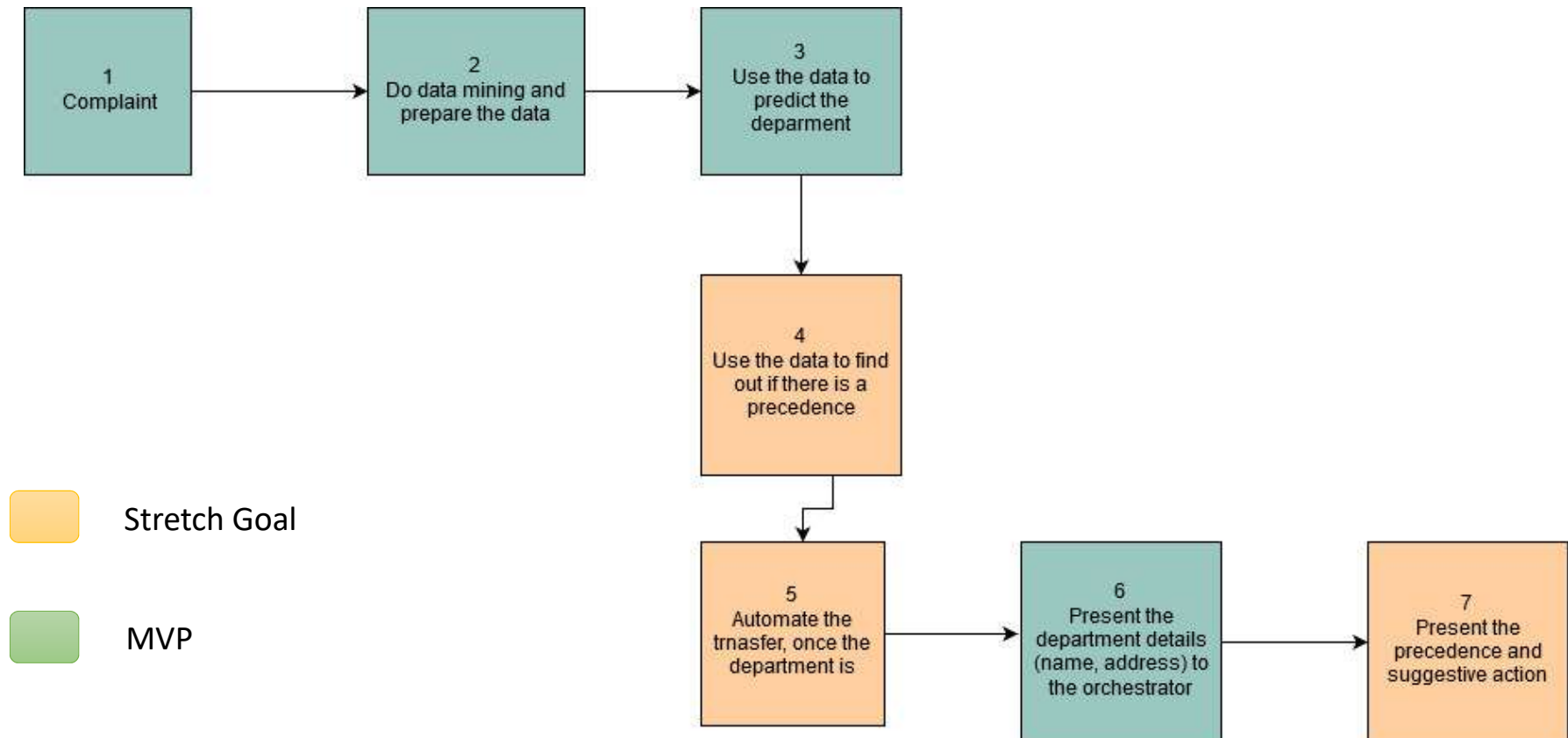
Complaints	Sum	% of Total
Total Receipts (01.01.2016 to 01.11.2019)	4779587	
Total Disposal (01.01.2016 to 01.11.2019)	3798755	79.48
Total Pending as on 01.11.2019	980832	20.52
Pending More Than 1 Year	599572	12.54
Pending Between 6 To 12 Months	100096	2.09
Pending Between 2 To 6 Months	118237	2.47
Pending Less Than 2 Months	162927	3.41

What is the problem we are targeting?

- One of the issue that we see is the transfer of complaint between multiple departments before being addressed. This takes a lot of time for the complaint to reach the right department/nodal officer to address it.
- Other issues we see are complaints sent back for more evidence or complaints disposed off, without taking any action.
- We also think that if nodal officers have information of previous remedial actions taken on similar complaints readily available, she/he might be able to take faster and affirmative action.

We are focusing on problem 1 as the MVP. Problem 2 and 3 will be treated as a stretch goal.

The workflow



The Demo

What to expect next

- Automate the transfer of request to concerned department
- Predict and suggest corrective actions
- Better modeling techniques – Ensembling , Neural Networks, Word Embeddings

Challenges Faced

- Subject_content is in many languages
- The data is heavily imbalanced
- Too Many departments. Need information on hierarchy of departments for accurate predictions.
- Text contains spelling mistakes and abbreviations are used.
- Feature matrix is too big and sparse.

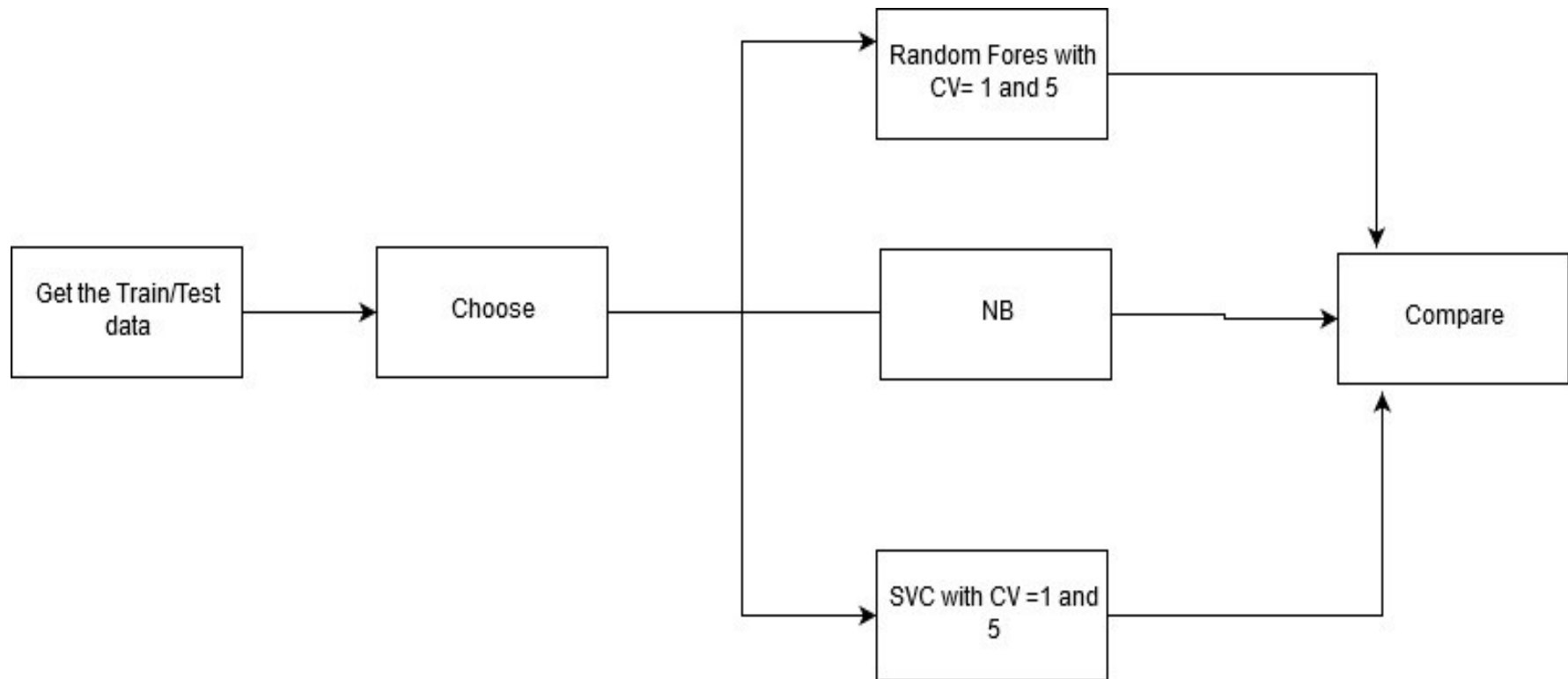
Data Analysis

- Total number of requests available in the dataset: from the **Public Grievance details in CPGRAMS along with feedback details**, we queried 2,60,000 records. From **Public Grievance movement details across the organization in CPGRAMS** we have queried 3,10,000 records.
- We assume that the right department is the department where the **action_name** is equal to: *'INTERIM REPLY TO COMPLAINANT', 'CASE DISPOSED OF', 'EXAMINED AT OUR LEVEL'*,
- Class distribution and final number of classes

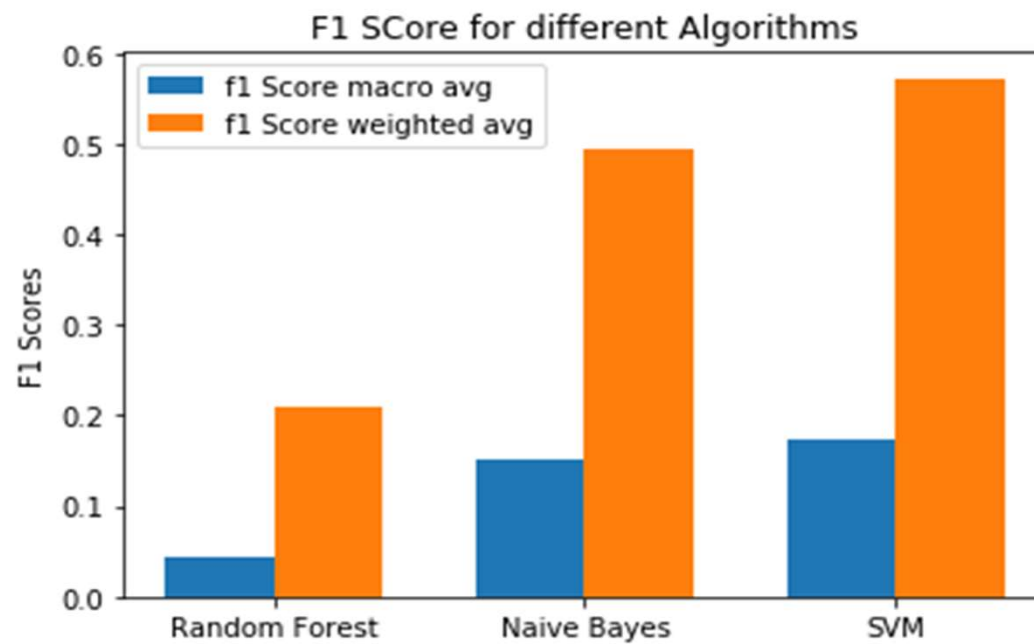
Details of final training data set

Data Set	No. of Records
Unique registration IDs in original data	222129
Registration IDs available in movement file	32890
Unique registration IDs where selected action_name available	29365
data size after removing 7500 rows from dominant class	21865
Final training set after removing 3000 (random) rows for blind test	18865

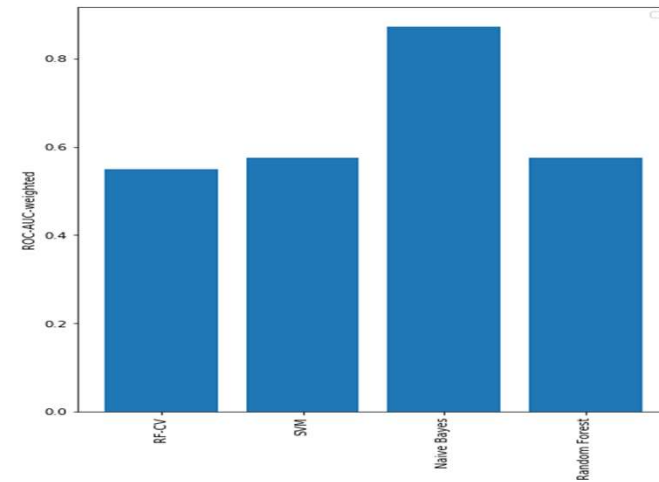
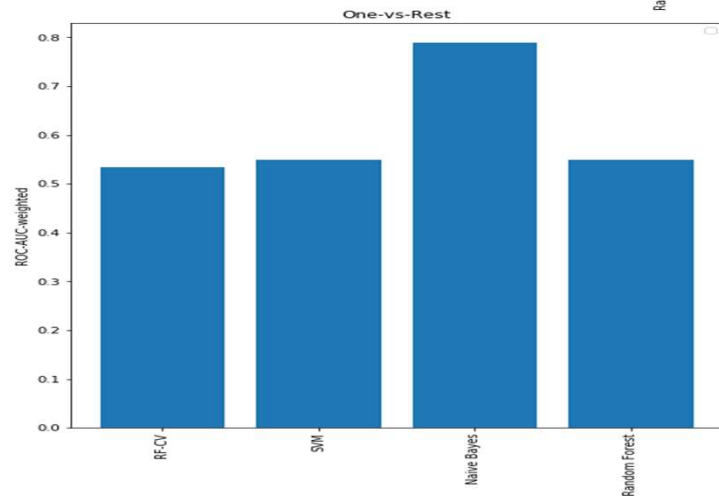
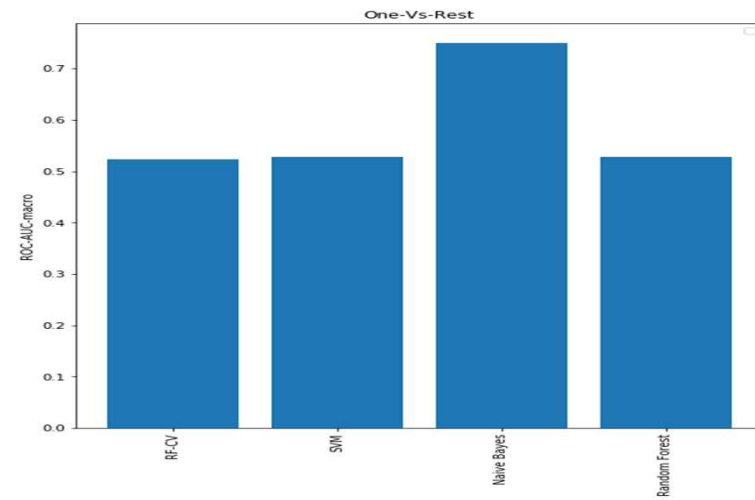
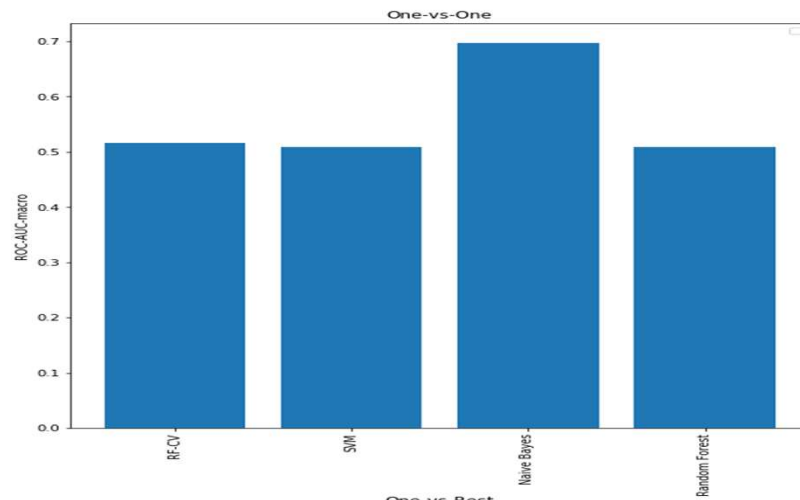
Model building



Metrics - 1



Metrics - 2



Thank You

For more details, please refer to the document “**DARPG-MVP and design.pdf**”

Appendix-1: Problem Discovery

- We did following research to understand the issues of the user and arrive at the solution proposed
- **Customer discovery:** We went through a lot of customer reviews/Quora discussion threads, for the pgportal site (and for the app also).
- The top problems which surfaced were:
 - Issues with logging in to the site and app
 - Time taken for the redressal.
 - Notification/updates to the user.
- **Personal experience:** We also consulted the friends and family to understand what was their experience of using the site. It was consistent with the above observations
- **Data Analysis:** Details of the analysis done on the data provided by the DARPG team, is given further in the document.