Paytm Problem Statement - 2 Hackfest 2019

Using technology to create a credit profile of e-commerce customers.

- By Team **Defaulting**,

IIT(ISM) Dhanbad

Problem Statement

The e-commerce companies find it very difficult to profit since there are lot of return backs and cancellations in e-commerce orders. Also even when orders are properly accepted the customer may want to pay by EMI. Thus their needs to be a system to create a credit and buying profile of a customer. This has to be done in a way that it does not create privacy issue for the customer. The developer should use data from customers and try to use technologies like big data/AI/statistical analysis to create a profile of customer without compromising his or her privacy. Then based on that the system should give delivery benefits and EMI options to the user. So a customer who has negative buying profile will not get Cash on Delivery option. Similarly customers with negative credit profile will not get EMI options.

Problem statement (in brief)

1. **Create credit profile: Assign a credit score to each customer. Only the customers above a certain threshold will be given EMI.**

2. **Create buying profile: Assign a buying score to each customer. Only the customers above a certain threshold will be given Cash on Delivery.**

IDEA

Data Mining + Deep Learning

IDEA

1. **A Data Mining system**
2. **A Deep Learning based classifier**

1. A **data mining system** which **feeds more data into the dataset we will create for classification**

   The data mining system will take input from several data stores periodically.
   Some examples:
   a. The data received from registrations can be used to match against several already downloaded databases of information **(some are mentioned above, namely, Caroll, MDR, Guidestar etc)** for probable matches. The matches found can thus provide more information about the newly registered users. Also, the **non-personally identifiable information** can be passed on to the third party APIs such as Dun & Bradstreet as well as Experian to find even more matches with a better confidence. T**he matches received from these registrations can be reconciled and the reconcile data can be used to improve the dataset which we will use for training the Deep Neural Network(Explained ahead) which will do certain predictions such as whether the users should be given EMI or not, or the category of products on which the offers should be given to the user.**

IDEA

1. **A Data Mining system**
2. **A Deep Learning based classifier**

b. The data from the databases that contain the details of orders of various customers will be harnessed by the data mining system to furnish the datasets which we will use for training the Deep Neural Network.

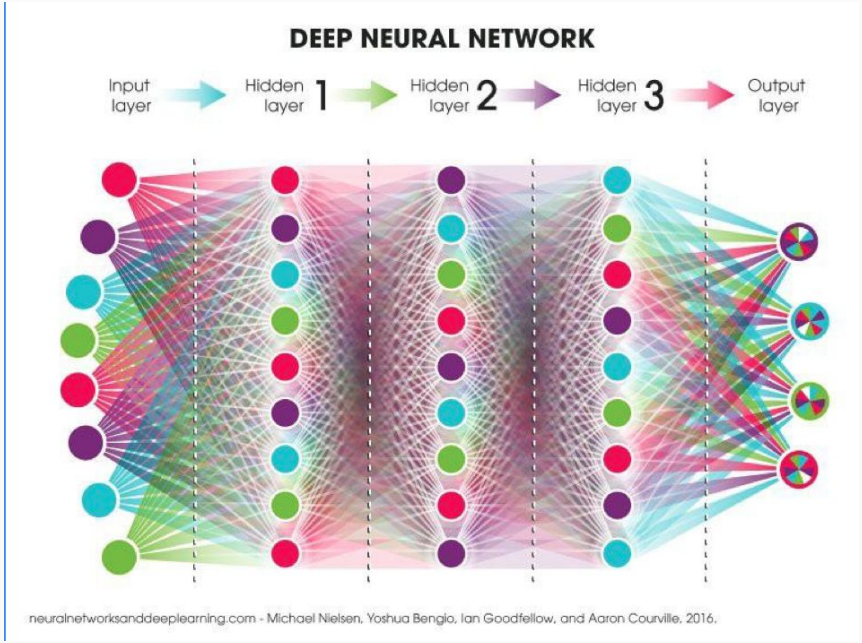
c. The data from the sales and advertising teams, such as whether the customer actually purchases a product after clicking on a previously sent advertisement will be harnessed to improve the predictions of the Deep Neural Network .

IDEA

1. A Data Mining system

2. A Deep Learning based classifier

2. A **Deep Learning Model(Deep Neural Network)** which will be used to predict certain traits about the customer based on the training data created using the
above mentioned data mining system. E.g. : Predicting related products based upon a customer's buying trends.

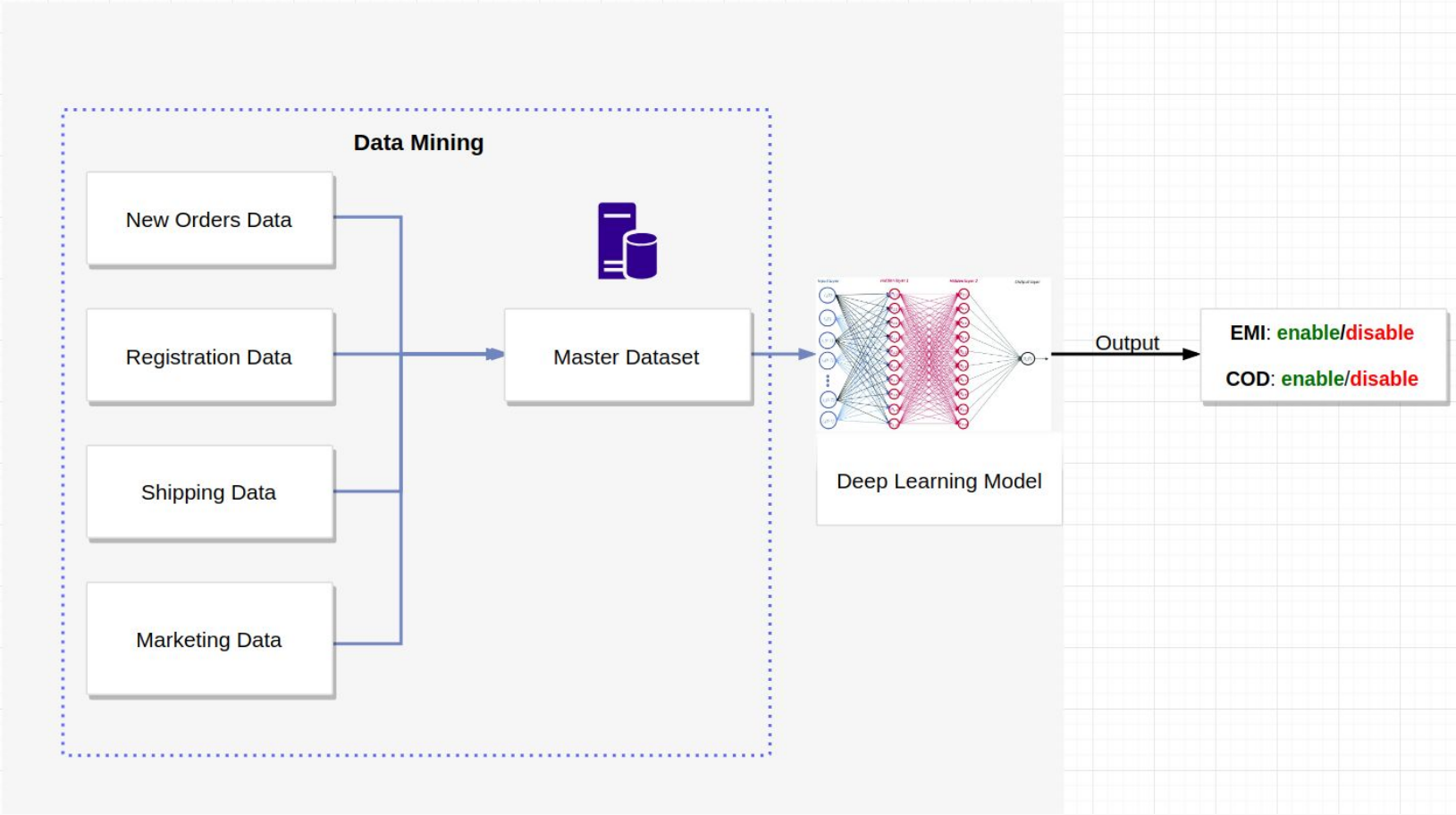


Master Data Table

```
SELECT * FROM `master_data`;
```

#	id	age	is_auto_billing	is_paytm_first	is_postpaid	postpaid_outstanding	orders_placed_in_6months	orders_placed_in_6months_via_epay	ord…
1	a4a0327847	45	1	0	0	0	5	4	1
2	a4a033ea47	40	1	0	1	1513	19	4	4

Architecture: Data mining and deep learning system



Highest in case of Neural Network (Multi Layer Perceptron)

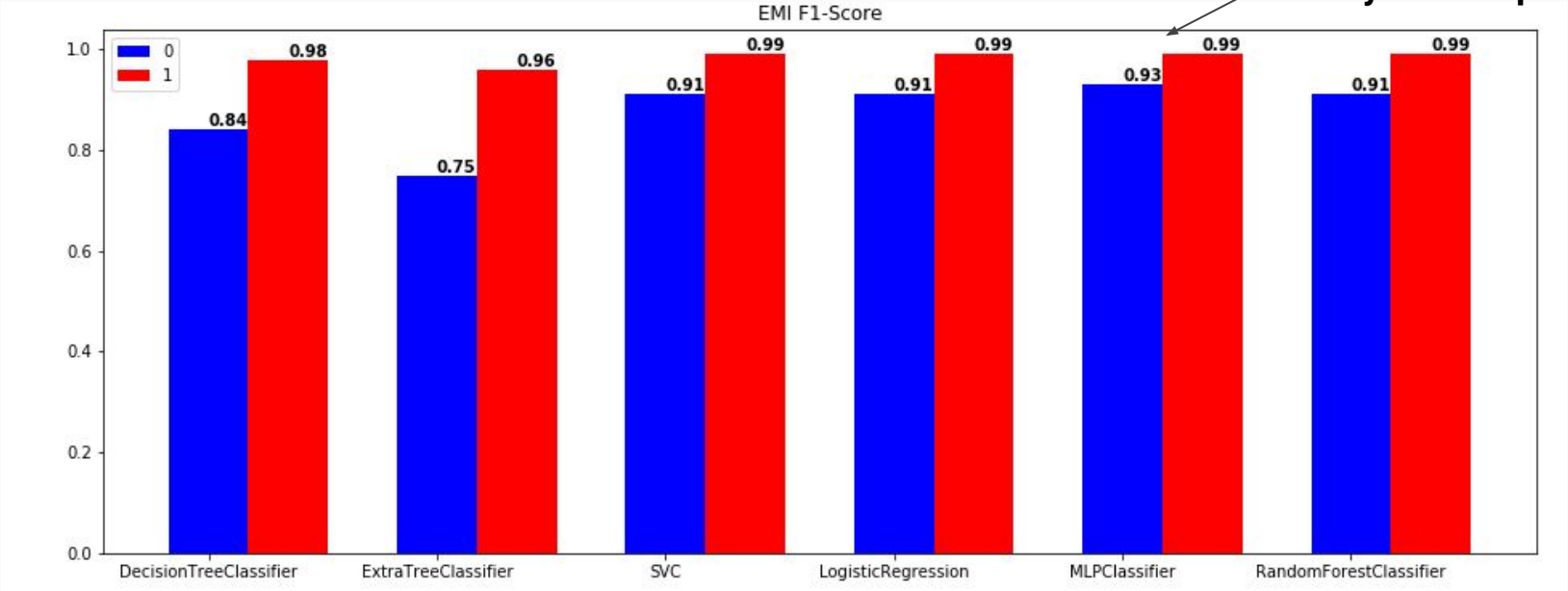


Fig: Comparison of F1 scores for EMI : Decision Trees, SVM, Logistic Regression, Multi Layer Perceptron, Extra Randomized Decision Trees, Random Forest

Highest in case of Neural Network (Multi Layer Perceptron)

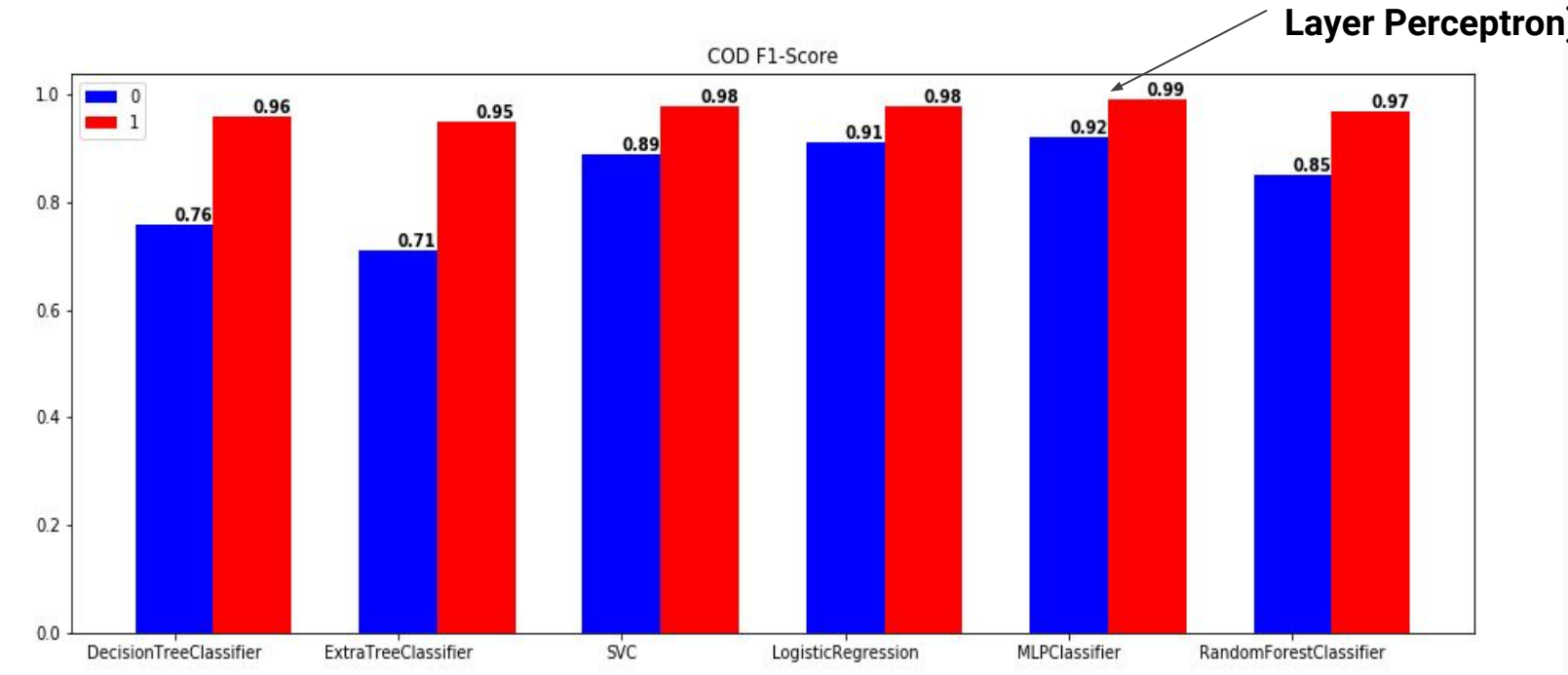
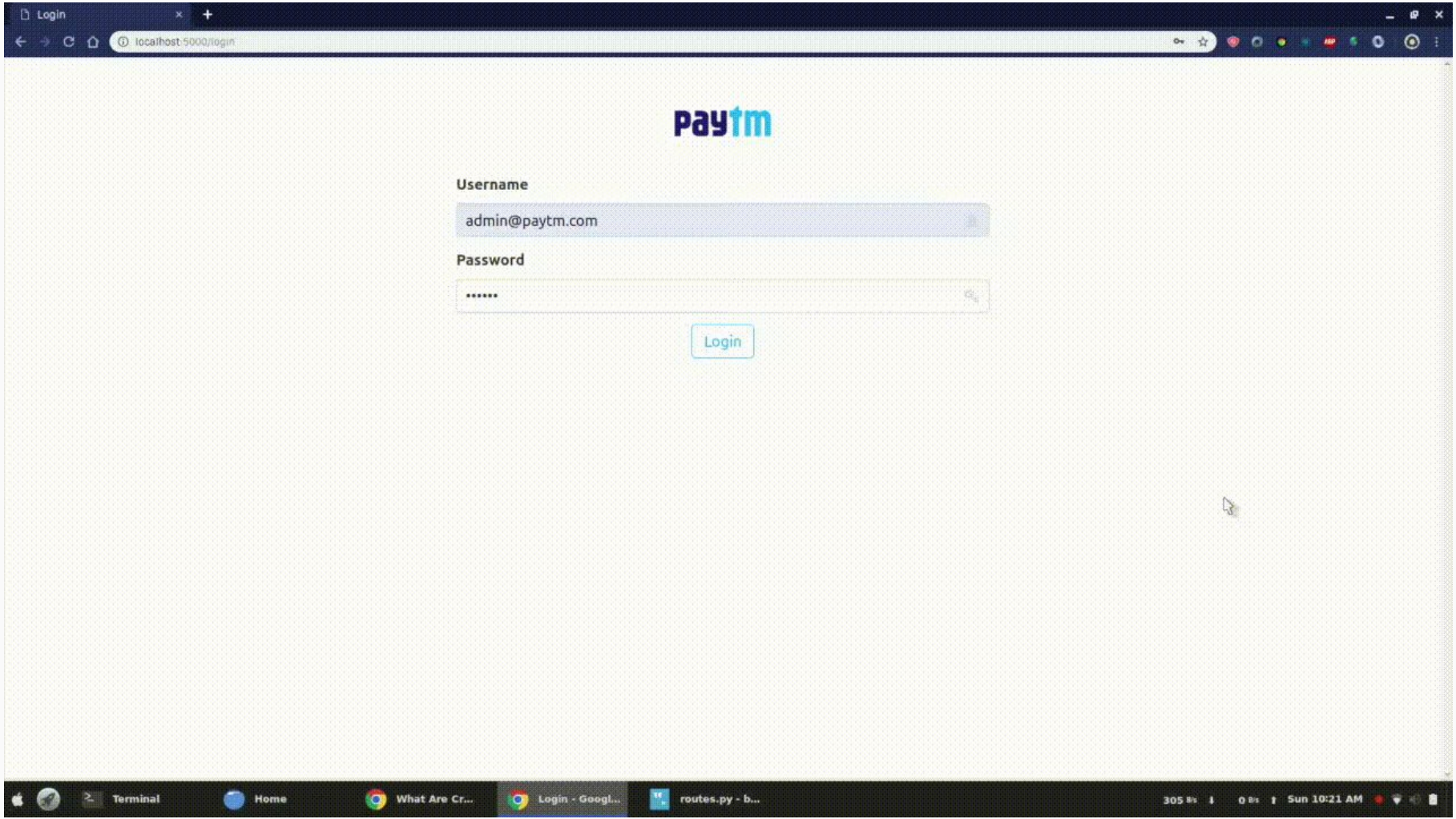
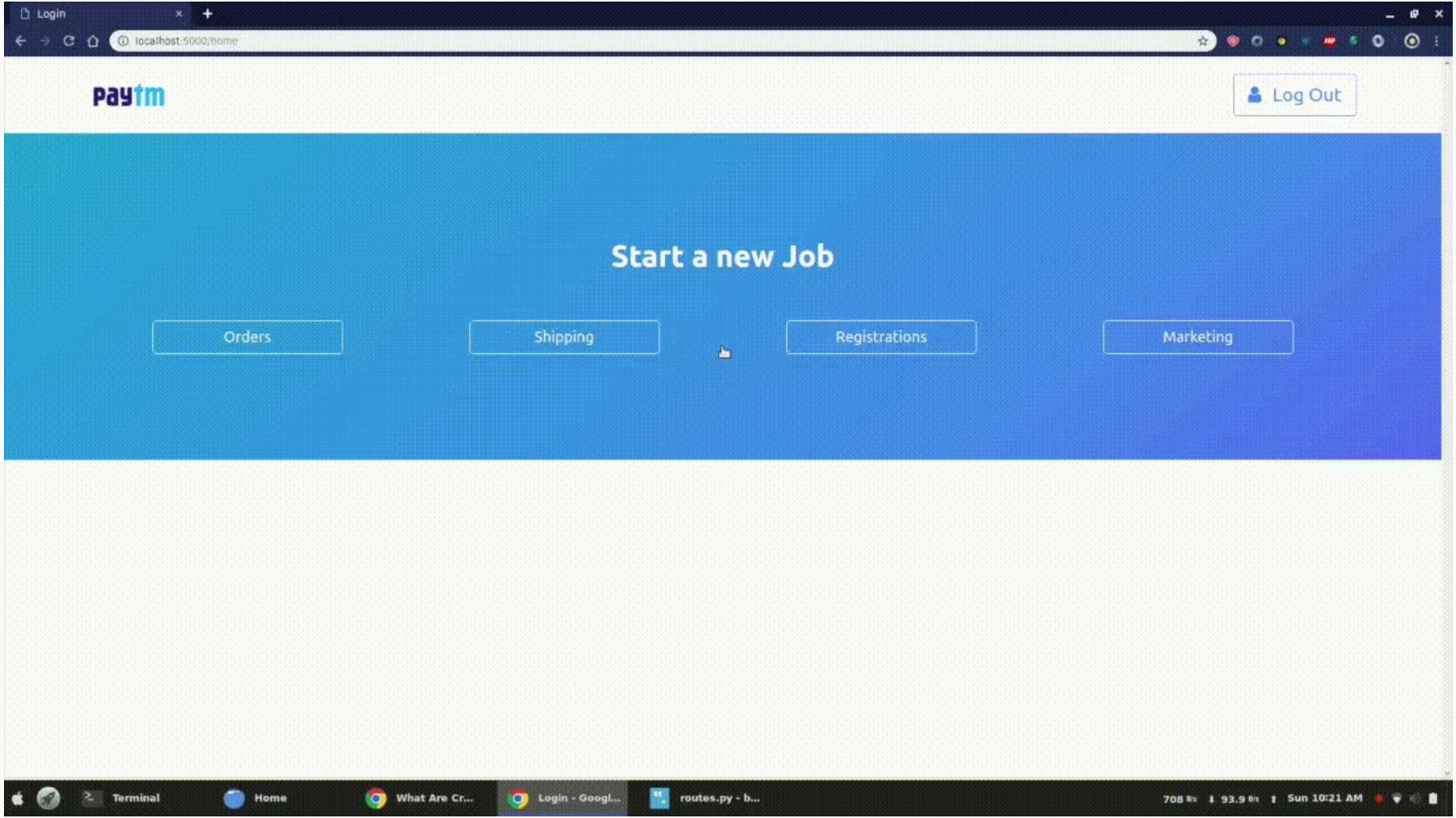


Fig: Comparison of F1 scores for COD : Decision Trees, SVM, Logistic Regression, Multi Layer Perceptron, Extra Randomized Decision Trees, Random Forest


Login
localhost:5000/login
paytm
Username
admin@paytm.com
Password
••••••
Login
Terminal
Home
What Are Cr...
Login - Googl...
routes.py - b...
305 B/s
0 B/s
Sun 10:21 AM


Login
localhost:5000/home
Paytm
Log Out
Start a new Job
Orders
Shipping
Registrations
Marketing
Terminal
Home
What Are Cr...
Login - Googl...
routes.py - b...
708 B/s
93.9 B/s
Sun 10:21 AM

Social Impact:

1. Highly beneficial for early age startups which offer Cash On Delivery, as predicting whom to give COD or not can prevent early age bankruptcy of budding startups.

2. Fair trade: Better credit score = Better market goodwill

Tech stack

- **Scheduler**: Bash scripts (cron jobs)
- **UI**: HTML, CSS, JS
- **Server Side**: Flask (Python based Microframework)
- **Data Storage**: Amazon Redshift, AWS S3 (Cloud file storage), AWS RDS
- **Other tools for testing purposes**: POSTMAN

Thank you

Team Defaulting