Progress Presentation-II

e-Yantra Summer Internship-2018
Machine Learning and It's Application

Swapnil Masurekar
Abhishek Sharma
Mentors: Rutuja, Suprabha

IIT Bombay

June 21, 2018

Overview of Project

Progress
Presentation-II

Swapnil
Masurekar
Abhishek Sharma
Mentors: Rutuja,
Suprabha

- Project Name : **Machine Learning and It's Application**
- Objective : To study machine learning and to work on its practical applications of Character Recognition, Automated Reply, Face Recognition along with documentation
- Deliverables :
  1. Learning ML and implement Applications
  2. Well commented code and documentation

Overview of Task

Progress
Presentation-II

Swapnil
Masurekar
Abhishek Sharma
Mentors: Rutuja,
Suprabha

Overview of
Project

Overview of Task

Task
Accomplished

Challenges Faced

Future Plans

Thank You

Task no.	Task	Status
Week 1		
1	Learning Basics of ML	Completed
2	Get familiar with Tensorflow(and python)	Completed
Week 2		
3	Character recognition using LR	Completed
4	Character recognition using neural network	Completed
Week 3 and Week 4		
5	Automated Reply System	Completed
Week 5		
6	Face Recognition	Ongoing
Week 6		
7	Documentation	Ongoing

Improvisations in Character Recognition

Progress
Presentation-II

Swapnil
Masurekar
Abhishek Sharma
Mentors: Rutuja,
Suprabha

Overview of
Project

Overview of Task

Task
Accomplished

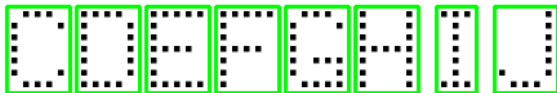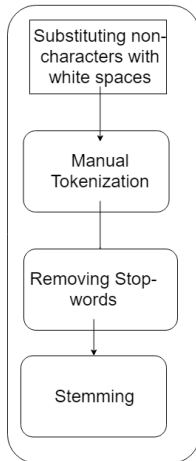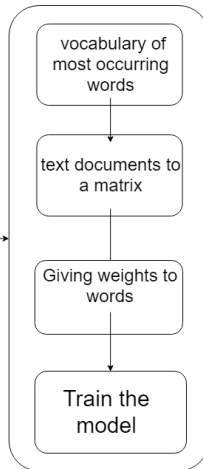Challenges Faced

Future Plans

Thank You

- Further expanding Dataset with Morphological transformations
- Sentence detection:

(x=284, y=1) ~ R:255 G:255 B:255

Results from Convolutional Neural Network: ====> TH1S IS A SENTENCE

- Dotted Character Recognition:

(x=312, y=62) ~ R:255 G:255 B:255

Results from Convolutional Neural Network: ====> CDEFGHIJ

Text Processing

Progress
Presentation-II

Swapnil
Masurekar
Abhishek Sharma
Mentors: Rutuja,
Suprabha

Overview of
Project

Overview of Task

Task
Accomplished

Challenges Faced

Future Plans

Thank You

Cleaning Text Data

Creating Bag of words model

Substituting non-characters with white spaces

↓

Manual Tokenization

↓

Removing Stop-words

↓

Stemming

vocabulary of most occurring words

↓

text documents to a matrix

↓

Giving weights to words

↓

Train the model

Gaussian Naive Bayes

- To label the mail in gmail inbox

Nearest Neighbor

- To generate the reply for the query

Figure: Automatic Mail Labeling using Gmail API

Automated Reply to E-mail query

Certificate ✉ Inbox ☆

SWAPNIL MASUREKAR
to learningml18
Hide details

From:	SWAPNIL MASUREKAR swapnil3597@gmail.com
To:	learningml18@gmail.com
Date:	Jun 15, 2018, 3:24 PM

View security details

Are we eligible for letter of participasion

learningml18@gmail.com
to me
Hide details

From:	learningml18@gmail.com
To:	swapnil3597@gmail.com
Date:	Jun 15, 2018, 3:24 PM

View security details

Check the portal for eligibility, if eligible you will receive certificates and goodies by end week of april

regarding power supply Inbox ☆

abhishek sharma
to learningml18
Hide details

From:	abhishek sharma ssshivalayabi@gmail.com
To:	learningml18@gmail.com
Date:	Jun 20, 2018, 3:16 PM

View security details

battery burnt due to heating so can we use power supply

learningml18@gmail.com
to me
Hide details

From:	learningml18@gmail.com
To:	ssshivalayabi@gmail.com
Date:	Jun 20, 2018, 3:17 PM

View security details

you can use auxilary power supply, but during finals you have to use battery

Challenges Faced

Progress
Presentation-II

Swapnil
Masurekar
Abhishek Sharma
Mentors: Rutuja,
Suprabha

Overview of
Project

Overview of Task

Task
Accomplished

Challenges Faced

Future Plans

Thank You

- Word Embeddings (word2vec)
- Finding best approach for finding most similar query to current query from Dataset:
  1. Sentence Similarity Based on Semantic Nets and Corpus Statistics
  2. Using Recurrent Neural Network for sequence classification
- Getting used to Gmail API
- Dotted Character Detection

- Optical Character-Recognition:
    - Detection of characters on ruled pages
    - Improvement in accuracy
- Face - Recognition

THANK YOU !!!