Yogesh Kumar

Email : yogesh.mcs17.du@gmail.com
Mobile : 8439142218

Education

- **Delhi University** — Delhi, India
  MSc Computer Science; 79.35% — *2017 - 2019*
- **Delhi University** — Delhi, India
  BSc Computer Science; 88.71% — *2014 - 2017*
- **PDDSVM, CBSE** — Vrindavan, India
  12th: 90.4%, 10th: 9.4 cgpa — *2009-2013*

Other Academic Qualification

- **UGC NET-JRF**: 99.96 Percentile, Rank-18
- **GATE**: 2019,2020 Qualified

Projects & Experience

- **MSc Reserach Project**
  Text Summarization
  - **Textual Entailment Based**: Implement and experiment with an existing algorithm based on Textual Entailment. Based on the results of the experiments create a new algorithm that uses graph properties to rank the sentences in the document which performs better for some of the ROUGE metrics.
  - **Graph Based**: Implement and experiment with two well-known Text Summarization algorithms TextRank and Corank that are based on the PageRank algorithm. We combine the TexRank and CoRank and evaluate the performance for the different contributions(linear addition of these two algorithms) to the new algorithm.
  - **LSA based**: Implement and experiment Latent Semantic Analysis based Text Summarization algorithms and study the effect of stop words as a preprocessing phase. Optimizes the algorithm by adding a step to check redundancy.
  - **Frequency based**: Implement an algorithm that uses word frequency to rank the sentences of a text document. Evaluate its performance and compare it with our other implemented algorithms that show word frequency is a good criterion to rank the sentence and can be included in other algorithms.
- **Plagiarism Checker**: This project aims to find copied documents from one another that are submitted as an assignment. Using K-means and hierarchical clustering with Cosine and Jaccard similarities documents clustered together to find similar documents.
- **Statistical Analysis of Open Source Projects**: Statistical analysis of Bug severity, priority, status, create time, resolve time etc. based on different software engineering metrics.
- **Binary Image Classifier**: A deep learning project based on logistic regression to classify images in different categories. This deep learning model trained with cat images that classify images with very high accuracy.
- **Salary Prediction**: A machine learning model using simple linear regression to predict the salaries of newly hired employees based on their year of experience.
- **Spent**: A chrome extension that keeps track of daily spending and notifies the user when the limit is reached. This project is implemented with Node.js, JavaScript, and HTML.
- **Predicting best profitable start-up to invest**: It is based on Multiple linear regression for predicting the best profitable start-up based on their R&D, location, number of employees, etc. using the backward elimination method with the highest p-values to eliminate redundant features.

Programming Skills

- **Languages**: C, C++, R, Python, HTML & CSS, Latex

Extra Curricular & Other Qualification

- Participated and was a rapporteur for the **"International Colloquium on Ethics and Governance of Autonomous AI Systems for a Better World"**, held on February 18 & 19 , 2019 organized by Center for Media Studies and supported by Department of Science & Technology, GOI.
- Internshala Student Partner
- Google India Udacity Challenge Scholar