Archit Kumar

archit.kumar@students.iiit.ac.in

9997295915

Muzaffarnagar , Uttar Pradesh, India

www.kaggle.com/archit29796

github.com/agarwal29796

SKILLS

Programming Langugages
C/C++ , Python , HTML , CSS , Javascript

Others
Machine Learning, System Design , MySQL , Linux

Soft Skills
Problem Solving , Critical Thinking , Project Management , Creativity .

Languages
English - Professional Proficiency , Hindi - Native or Bilingual Proficiency

KAGGLE

IEEE-CIS FRAUD DETECTION
rank - top 21%

GENERATIVE IMAGES
rank - top 18%

INTERESTS

Machine Learning , System Design , Data Science , Deep Learning

ACHIEVEMENTS

Air 260 - GATE CSE 2018

EDUCATION

MTech in Computer Science
International Institute of Information technology , Hyderabad

07/2018 – Present *CGPA - 8.23*

Courses
- Database systems
- statical method in artificial Intelligence
- Scripting and computer environment
- Operating system
- Information Retrival and Extraction
- Computer vision

BTech in Electronics and Comm. Engineering
Bundelkhand Institute of Engineering and Technology

07/2014 – 05/2018 *Percentage - 77.75%*

PROJECTS

Wiki Search Engine (08/2019 – 09/2019)
- The project aimed to implement a text based search engine on 70 GB of Wikipedia dump .
- We implemented an inverted index (sorted by tf-idf score) based search engine . Which resulted in ~8GB size of invered index and achieved average searching time of 500 milliseconds .
- Key Concept : Infromation retrival , SAX Parser , Stemming , Python , WikiPedia dump , Inverted Indexing .

Recognition of Handwritten Words and Lines using a CNN-LSTM Network (03/2019 – 04/2019)
- The Project aimed to implement a neural network architecture to detect handwritten words for IAM dataset.
- Experimented different augmentation techniques (ex: Affine transformations) for image generations. Augmentation was performed to make neural network invariant to the changes in individual handwriting but highly variant for the handwriting of two different people.
- Key Concept : CNN-LSTM neural network , CTC loss , Augmentation .

Http Server in cpp using thread pool (10/2018 – 11/2108)
- The project aimed to implement the core features of http server (get and head methods) in c++ using pthreads library .
- We Implemented a multithreaded scalable server and presented different analysis of the performance of the server for thread counts , cpu cores , numbers of tasks .
- Key Concept : Threadpool , Threads Synchronization , C++ , Pthread Library , Socket Programming .

SemEval19: OffensEval (09/2019 – Present)
- The project is multiclass hierarchical in nature classifcation . Classification categories includes Offensive language identification , Automatic categorization of offense types and Offense target identification .
- Experimented with various classifcation algorithms . Analysed performance of different model for different input text (tweets) representation like bag of words , tf-idf .
- Key Concept : Hierarchical multiclass classification , Text representations (Bag of Words , TF-IDF) , Python , Char-CNN neural network , Feature Engineering ,

Malicious web pages detection (03/2019 – 04/2019)
- Malicious URL Detection using Machine Learning: Malicious URL, a.k.a. malicious website is a common and serious threat to cybersecurity. Malicious URLs host unsolicited content (spam, phishing, drive-by exploits, etc.) and lure unsuspecting users to become victims of scams (monetary loss, theft of private information, and malware installation).
- We extracted and searched different features (ex: URL length, DNS response time, etc.) from a given URL and experimented with various machine learning algorithm that could categorize benign and malicious web pages separately .
- Key Concept : Feature Engineering , Honey Pots , Selenium Browser , Javascript Parsing , Python , Precision/Recall ,