

# ANUSHKA JAIN

(470)-807-5207 | ajain32@student.gsu.edu | <https://www.linkedin.com/in/anushkajain-21> | [github.com/anushkajain21](https://github.com/anushkajain21)

---

## EDUCATION:

### Master of Science in Data Science and Analytics

Aug 2021-Dec 2022

Georgia State University, J. Mack Robinson College of Business

Relevant Coursework: Machine Learning, Deep Learning, Natural Language Processing, Data Warehousing, Data Visualization, Data Analysis, Predictive Analytics

### Bachelor of Engineering, Computer Science

Aug 2014-May 2018

Acropolis Institute of Technology and Research - Indore, India

---

## SKILLS:

- Skills: Python, R, Machine Learning, Clustering, PCA, Regression, Classification, Recommender Systems, Neural Networks, Data Mining, Data Wrangling, Spark
  - Data Skills: Hadoop, SQL, MongoDB, PostgreSQL, Tableau, MS Excel
  - Libraries: Pandas, Numpy, Matplotlib, Scikit Learn, Seaborn, OpenCV, TensorFlow, PyTorch, SpaCy, Selenium, Gensim
  - Natural Language Processing: Bag of Words, Topic Modeling, TF-IDF, Text Analysis, Sentiment Analysis, Word2vec
- 

## WORK EXPERIENCE:

### Graduate Research Assistant– Institute for Insight

Aug 2021 – Present

- Experimented NLP Techniques to build a scalable and efficient NER pipeline to identify extremist content from social media
- Implemented models like LDA and word2vec for topic modeling and SpaCy for context mapping for both extremist and non-extremist data

### Assistant System Engineer – Tata Consultancy Services

Aug 2018 – Jul 2021

- Improved existing system to gather generated data using Pyspark to verify and perform time synchronization and ended up saving 12% of time
  - Programmed a solution to ingest data in Hadoop File System and transformed it to match the automotive standards using PySpark for handling data efficiently
  - Engineered a module using Pyspark that extracted ingested data based upon user requirement for 50+ users
  - Developed algorithms using python to solve client requirements for Path Prediction, Event Detection and Object Detection
  - Designed a component for generation of graphs using visualization libraries and increased business by 2%
- 

## PROJECTS:

### Prediction on Kickstarter Data

Nov 2021 – Dec 2021

- Performed Data Preprocessing using Pandas and EDA using matplotlib, seaborn and missingno to create 5+ visualizations
- Evaluated the performance of 8 implemented classification models using scikit-learn resulting in 91% accuracy

### EDA on US Police Shootings Data

Nov 2021 – Dec 2021

- Determined the parameters responsible for shooting from 15+ graphs generated using visualization libraries like matplotlib, seaborn and folium

### Web Scraping and Analysis of Estimize.com

Nov 2021 – Dec 2021

- Streamlined the Web Scraping process for Estimize.com using selenium and BeautifulSoup to store data into mysql database for analysis
- Used scikit-learn to apply Regression model to predict the accuracy score of 100+ analysts

### intelliShopping – The Smarter Way to Shop (Patent Pending)

Dec 2016 – Aug 2018

- Orchestrated Internet of Things based solution to resolve cumbersome billing system in a supermarket to bring down consumer waiting time from 25 minutes to under 2 minutes
  - Designed and constructed an analytical dashboard to generate consumer behavioral patterns for 10+ consumers
- 

## ACCOMPLISHMENTS:

- Received Merit Based Scholarship of 15000 USD from Robinson College of Business in 2021
- Recognized as one of the TOP Startups by Confederation of Indian Industry: i3 Innovation conclave and Top 25 Startups for Indo-Canada Acceleration program by All India Council for Technical Education in 2018
- Earned Runner-up in EnginX IoT Challenge 2017 held worldwide by Tata Consultancy Services