# Praneeth Reddy Soddum

[soddumpraneethreddy95@gmail.com](mailto:soddumpraneethreddy95@gmail.com%20) | 2052535752| <https://www.linkedin.com/in/praneeth-soddum-9958021b3>

<https://github.com/pr365>

A Data Analyst with 3 years experience in Data Analytics and expertise in building machine learning models (supervised and Unsupervised learning) and NLP models. Skilled in handling multiple tasks from Data processing, Drawing insights from data by using statistical concepts, Building models and validating the model. Currently a Teaching Assistant and working for a research with a professor by building visualizations and dashboards.

# RELEVANT EXPERIENCE

**Infosys Limited**

**Data Analyst Jan2017 – Dec 2019**

* Handling full stack development and integration tasks, data extraction and processing from large databases and building predictive models.
* Developing Predictive models to analyze the claim strategies in insurance and healthcare domain to make profitable decisions.
* Recognizing the patterns and drawing actionable insights from the data by using Data mining techniques and statistical methods.
* Executes data retrieval and reporting using SAS and other data management tools for data analytics, model building, and model monitoring
* Built ETL pipelines and performed ETL operations over business data including extracting from different sources, applying transformations, and loading in DB2. Resolving critical job abends in online, batch and perform root cause analysis.
* Building predictive models using Regression, Ensemble techniques to see the trend in claim categories based on business requirement.
* Create dashboards and reports that present meaningful data insights and business opportunities to technical and non-technical stakeholders.
* Work with business leaders to define key metrics and build reporting to monitor and understand performance along those metrics
* Conduct in-depth data analyses that lead to actionable insights, owning the entire process from ideation to execution
* Develop data models in our data warehouse that enable performant, intuitive analyses

# Infosys Limited

**Data Analyst Intern Jan 2017 – May 2017**

* Report actionable, statistical and analytical insights to executives for effective decision making.
* Maintained large database and used professional statistical methods to collect, analyze and interpret results by visualization tools.
* Write python scripts to automate the daily tasks, data extraction and data cleaning by executing the CRON jobs

# EDUCATION

**Master of Science in Data Science (GPA 4.0) Jan 2020- Aug 2021**

University of Alabama at Birmingham

# TECHNICAL SKILLS

**Programming Languages/ Frameworks:** Python, PYSPARK, Java, AWS, HTML, XML, CSS, Node.JS,D3.JS,COBOL,FLASK,Django

**Database: SQL:**DB2,PostgreSQL,MYSQL, **NOSQL**: MongoDB, Apache Cassandra

**Tools:** Anaconda, Tableau, GitHub, Visual Studio, Power BI,SAS,Excel(pivot table,vlookup),ETL(Informatica)

**Coursework:** Machine Learning, Data Mining, Deep Learning, Quantitative and Qualitative for Managers, Advance Database systems, Statistical Analysis, Cloud Computing

# CERTIFICATIONS

**AWS Solutions Architect Certification July 2021 – July 2024**

**Skills:** EC2, VPC, RDS, S3, Lambda, ECS, IAM, Cloud Front, Snow Flake, DynamoDB, Route53, Redshift, Sage maker

# Data Analysis and Machine Learning with Python May 2020 – till date

**Skills:** Regression, Bagging (Random Forest, Decision Tree), Boosting (XG Boost, Cat Boost), Clustering, Neural Networks (ANN,CNN,RNN)

# RELEVANT PROJECTS

**Customer Segmentation and Market Basket Analysis Data Mining**

* Implementation customer segmentation using Clustering based on RFM metric.
* Used Association rule mining algorithms like FP-growth, ECLAT and Apriori for frequent item set mining

# Machine Translator using GRU, RNN using Embedding, Bidirectional RNN and Deep RNN NLP(Sequence2Sequence)

* Implemented simple RNN using GRU,RNN using Embedding, Bidirectional RNN to convert text from English to French, Finally stacked all these algorithms and constructed Deep RNN which has achieved highest accuracy.

# Product review Sentiment Analysis using MNB, Logistic Regression, LSTM and deployed best models using Flask Deep Learning

* Implemented Multinomial Naïve Bayes Model using Count vectorizer and TF-IDF vectorizer as feature extraction technique.
* Implemented a pipeline model which has TF-IDF and logistic regression using Grid search to get best parameters.
* Implemented LSTM model using tokenizer and word2Vec embedding with hyperparameter tuning.

# Object Detection using Mask RCNN – Transfer Learning Machine Learning

* Used transfer learning to detect the nuts and threads from an image. Coco dataset model weights were used for training the model.

# Interative COVID 19 Tree Map and US map visualization Data Visualization

* Developed an Interactive week wise Covid-19 Tree Map Visualization. A lollypop chart will be displayed to see continent wise covid cases.
* Click on to see the visualization - <https://pr365.github.io/covid-treemap-/>