# EDUCATION

SAGAR VARMA SAMANTHAPUDI

varmasagar7799@gmail.com|+1 (513) 837 - 5273 linkedin.com/in/sagarvarma | [github.com/sagarvarma](https://github.com/sagarsamanthapudi)

**Indiana University-Purdue University (IUPUI),** Indianapolis, US ***August 2021 – May 2023***

Master of Science in Applied Data Science GPA: 3.95/4

***Courses***: Database Design, Machine Learning, Cloud Computing, Data Visualization, Probability & Statistics, Data Analytics

***Activities*:** Teaching Assistant for Object Oriented Programming & Software Development in MVC (Class Size: 35)

**GITAM University,** Vizag, India ***July 2013 – April 2017***

Bachelor of Technology in Mechanical Engineering GPA: 3.5/4

# TECHNICAL SKILLS

**Programming :** Python, R, C++, SQL

**Techniques :** Machine learning, Deep learning, Linear & Logistic Regression, KPSS Test, Hypothesis Testing, Apache Spark, Moving Average Modelling, K-means Clustering, Natural Language Processing (NLP)

**Data Visualization :** Tableau, Power BI, Microsoft Excel

**Libraries :** NumPy, Pandas, Scikit-Learn, Tensor Flow, Feature Engine, ggplot, Seaborn, Matplotlib, Keras, OpenCV

**Databases :** MySQL, Oracle, PostgreSQL, Amazon S3, Amazon Dynamo-DB, Mongo DB, Spark SQL

**IDE - Tools :** Git, Visual Studio, PyCharm, AWS, GCP, Microsoft Office

**Soft Skills** : Leadership, Team Management, Agile Practitioner, Creative & Confident Speaker

# EXPERIENCE

**The Polis Center, School of Informatics & Computing, IUPUI |** *Data Science Intern* ***February 2022 – May 2023***

* Developed ETL Pipelines for data processing from multiple sources to SQL Server using SSIS packages and load into data lake for application of other polis projects.
* Created online assets data dashboard by transforming raw data into actionable information using Power BI for strategic planning and evaluation by Indiana State Government.
* Enhancing SAVI community systems data-pipeline using predictive modelling and integrating with geo-spatial maps using Arc-GIS to analyze and visualize the data about local communities.

**School of Informatics & Computing, IUPUI |** *Data Science Research Assistant* ***August 2021 – January 2022***

* Built an automated Machine Learning data pipeline for extracting overview of datasets using supervised NLP technique Bert for entity recognition with an accuracy **81.4 %**
* Applied Sherlock for detecting semantic types and mapping the data column to dictionary with support weighted F1 Score **82%**

**Tech Mahindra Ltd, Delhi India |** *Software Engineer* ***April 2018 – July 2021***

* Lead the development team for the project of two e-commerce web application Cenveo & Discount Labels (Subsidiary of GE).
* Implemented logistic regression model for order success rate prediction using sklearn and RF regressor for cost prediction gained **78%**

accuracy.

* Introduced a new feature provides business team an elastic search functionality to optimize the product recommendation for customers.

**Global Logic, Hyderabad India |** *Associate Analyst Intern* ***June 2017 – March 2018***

* Developed a tool for automatically labeling the road markers for preparing road-marker dataset for self-driving cars.
* Implemented Marker Based Watershed Segmentation and Harris Corner Detection algorithm for detecting road markers.

# PROJECTS

## Real Estate Application

* Implemented support vector machine and logistic regression algorithms to predict hotel availability with an accuracy of 89% on test data.
* Performed data wrangling and data visualization for exploratory data analysis and dealt with categorical and unbalanced data.

**Dog Breed Classification – Udacity Nano Degree AI Programming** *(Python, Deep Learning, Keras, Pandas, Numpy, Matplotlib)*

* Predicted dog breed labels and classified into categories of ImageNet dataset using CNN implemented in Keras with an accuracy of **72%**
* Executed Transfer Learning CNN by choosing InceptionV3 as the network which provided the features for training layers.
* Employed Data Augmentation to prevent overfitting add more layers to make model more complex in prediction.

## Movie Library Application

* Developed a movie library application utilized dynamic programming principles and with rich user experience.
* Demonstrated classic programming algorithm merge sort to classify and display movies by genre.

## Guessing Number Puzzle Game

* Developed a number guessing game applied Binary Search Algorithm tasked for finding a number within a range.

# AWARDS AND ACHIEVEMENTS

* Udacity Bertelsmann Business Analytics Nano Degree Technology Scholar 2022
* Remote sensing data classification Deep Learning researcher Indian Space Research Organization, Dehradun 2021
* Elite Tech-mighty for the batch of 2018 among 1384 associates
* Won 1st prize in “Internet of Things”, a coding competition in the Techno Fest 2k16 of IIT Bombay