SURYA SRIKAR SIRIGINEEDI

+1 (786) 451-0582 | suryasrikar.s@gmail.com | https://www.linkedin.com/in/suryasrikar/ |https://github.com/SuryaSrikar

SUMMARY

Data Scientist with 3+ years of experience in project work in predictive modeling, data processing, ETL, application development to solve challenging business problems with technology. Beat 250+ participants in **shell hacks 2019**, by building models that provided best solution for the challenge. And a team player with strong interpersonal skills and communication skills. Involved in open source community and passionate about deep learning.

TECHNICAL SKILLS

- **Programming Language:** Python, T-SQL, C#, C++, SQL, Scala, C, Py Spark, MATLAB.
- **Machine Learning Framework:** Scikit Learn, Tensor Flow, FB prophet, Keras, Py torch.
- **Big Data Tools:** Spark, Kafka, Neo4j, and Azure Data Studio.
- **Database:** MS SQL, My SQL, Big Query, HDFS, MS SQL, Mango DB, MS Access 2007.
- **Data Visualization:** matplotlib, bokeh, seaborn, py plot, Tableau, SAP Crystal Report, Power BI, MS Chart, SAS.
- **Cloud:** Microsoft Azure, AWS, GCP.
- Strong knowledge on OOP Concepts and design patterns.
- **Computer Vision:** AForge, Data Augmentation, Object Detection, Optical Flow.
- **Natural Language Processing:** Lemmatization, Stemming, Named Entity Recognition, TF-IDF.
- **Algorithms:** Regression, Decision Trees, Clustering, Association Rules, K-Nearest Neighbors, Neural Nets, SVM, Ensemble.

RESEARCH PUBLICATION:

2020 The 4th International Conference on Compute and Data Analysis: "*Learning-based models to detect runtime phishing activities using URLs"* by **Surya Srikar Sirigineedi**, Jayesh Soni, Himanshu Upadhyay.
2020 GIS Conference *"Internet enabled remotely controlled architecture to release water from storage units"* by Vivek Verma, Arturo S. Leon, Dogukan Ozecik, Linlong Bian, **Surya Srikar Sirigineedi**. (Accepted)

PROFESSIONAL WORK EXPERIENCE (3+ Years)

APPLIED DATA SCIENTIST – FIU APPLIED RESEARCH CENTER, USA — Jan 19 – Till Date

Exposure: Anomaly Detection, Tensor Flow, HDFS, Spark, Microsoft Azure, Kubernetes, NLTK, CNN, Kafka.

- Collaborated with team members to build a **Host Based Intrusion Detection System** that identifies the compromised VM machine with system calls and process memory behavior, achieved 89% accuracy using 1 class SVM and LSTM.
- Read paper and deliver insights on neural networks (CNN, AE, VAE, GAN) supervised and unsupervised models.
- Took ownership as architect and developed a wound monitoring system using wearable sensor which helps in identifying infections during wound healing process by performing Health Care Analytics on edge devices.

MACHINE LEARNING INTERN - CBS INTERACTIVE, USA — May 19 – July 19

Exposure: Time series Analysis, SARIMA, FB Prophet, AWS, Google Big Query, My SQL, Docker, Grafana.

- Objective: Utilize historical data of CBS Sports site traffic to build the prediction model which can forecast server capacity required, this model was used to by Dev Ops to plan the server capacity required for the day.
- Designed and developed the ETL process to Extract translates and load data from Big Query GCP to My SQL on AWS.
- Performed quantitative analysis on Time series data, extracted unique user visitors per day, visit duration, page views from peta bytes of data from Google Big Query and generated hypotheses to site traffic.
- Automated the pipeline for reporting and developed user interface with reproducible method using Grafana and

deployed developed model in production on AWS cloud and exposed the model as API using flask and Grafana.

- Installed and supported entire **CI/CD** pipeline on AWS Cloud for the application.

Sr. MACHINE LEARNING ENGINEER - GGK Technologies **Mar 17 – Aug 18**

Exposure: MS SQL, SSIS, Microsoft Azure, Angular2, Web API, Dot Net, WPF, Power BI.

- Developed SQL Complex Queries for ETL mappings validation, ad- hoc reports and quantitative & qualitative analysis with business analysts' team of **Inovalon** USA health care company. Where I used to deal with **1M+ patients** EHR.
- Collaborated with different team members to designed **cloud pipeline** to process events from **AT&T** and **Sky Bitz** IOT GPS devices of 30k trucks of **Schneider Logistics** USA in real time cloud infrastructure on **Azure Cloud**.

MACHINE LEARNING ENGINEER - ZEN Technologies **Dec 14 – Mar 17**

Exposure: Dot Net, C#, My SQL, Distributed Systems, Power BI, SAP Crystal Report, Image Processing, MFC (VC++).

- Created dynamic reports using SAP Crystal Reports to get insight about trainee progress and provided the intelligent feedback mechanism which identifies the areas to improve to fast up the training process.
- Developed application in MFC (VC++) to perform monochrome camera **calibration**. Built an application which auto detects the moving objects in video using AForge framework, which has reduced a lot of **manual effort**.

CERTIFICATION

Data Science with Neo4j, Introduction to Neo4j, Data Science Essential Training, Bharat Sanchar Nigam Limited (BSNL) in plant training.

EDUCATION

Master of Science in Computer Engineering **Aug 2018 - May 2020**

Florida International University, Miami, Florida **GPA: 3.8**

Bachelor of Technology in Electrical & Communication Engineering **Aug 2011 - May 2015**

Jawaharlal Nehru Technological University, Hyderabad, India **GPA: 3.7**

EXTRA CURRICULAR

AWARDS: Runner up for MITRE Shell Hacks Hackathon (out of 250+ participants), SGA Graduate Scholarship.
POSITION: Treasure for FIU Indian Cultural Club. Worked as Vice President at IETE Student branch.
BUSINESS: Built my own computer service center business through my college days.

PERSONAL PROJECTS

MOVIE RECOMMENDATION SYSTEM: Python, NLTK, Surprise, Cosine Similarity

Recommender System is a system that seeks to predict or filter preferences according to the user's choices. The analysis involved text data preprocessing and feature engineering (tf-idf) using NLP techniques.

NETWORK INTRUSION DETECTION SYSTEM: Auto Encoder Decoder, CNN, Anomaly Detection, Research

Detecting the malicious activity in the network using the network encrypted payload, by converting the payload into an image format and then using Auto Encoder Decoder Neural Networks able to classify the payload into benign or not.

PHISHING URL DETECTOR: NLP, Sci kit Learn, Logistic Regression, KNN, Deep Learning, Python

Applied feature engineering technique on the phishing URL to extract the features from given URL like URL length, Special characters, top level domain, second level domain, NLP etc. Used extracted features to classify the URL is malicious or not.

TELECOM CUSTOMER CHURN: Python, KNN, Random Forest, GB, XGBoost, Ensemble

Build the machine learning model using KNN, Random Forest, GB, XG Boost after preprocessing the given dataset using label encoder and One-Hot Encoder to predict customer attrition with 89% accuracy.