Brinda Somasundaram

brinz99@gmail.com /*(862) 485 9331*

Summary of Qualification and Objective:

- **Highly analytical and process-oriented data analyst** with in depth knowledge of database systems; research methodologies; and big data capture, curation, manipulation and visualization.
- Vast knowledge in science and health; highly appreciated projects undertaken during ***Masters in Health Informatics and Data Analytics*;** specialization in data science and analytics including statistics, data mining, and predictive analytics.

Core Competencies:

- Proficient in using **UML** for Business Process Modeling, Use Cases, Activity Diagrams, Sequence Diagrams, Data Flow Diagrams, Collaboration Diagrams, Class Diagrams, and Wireframe Prototypes
- Big Data ecosystem, **Hadoop**, **Spark** and Spark SQL using HDFS and RDD concepts. Setting up a **Data Analytics Platform** primarily for **Data Mining** on Health data.
- **RDBMS** databases: **MS SQL Server**, **Oracle, PostgreSQL**.
- **No-SQL** databases: **Influx** DB a time series database, **Neo4j** the most popular graph database and **Mumps** a general-purpose programming language that supports a novel, native, hierarchical database facility.
- Strong knowledge and understanding of **Data Modelling** concepts, including UML diagrams.
- Data Analysis using **Python** Programming Language (libraries used: NumPy, SciPy, Matpotlib, Scikit etc.) and **R-Programming** Language for advanced statistical analysis.
- Proficient in Data Analytic software - **SAS Statistical tools, Stata** and **Weka**- Data Mining software.
- Data visualization tools such as **Tableau**, and **Microsoft Power BI**.
- Confident, articulate and **professional presentation**, **writing** and **speaking** skills.

Experience:

REI Systems, -Health and Human Services(HHS), Health Resources and Service Administration(HRSA), Rockville, Aug 2017 - Present

More than 20% of Federal spending is done in the form of grants. State and local governments, nonprofits, private organizations, and individuals receive them. REI's grants management services make that happen.

Responsibilities:

- Data Mapping and Analysis.
- Preparing reports and dashboards using excel, **SQL and Tableau** on various **clinical measures** of various the **Health Centers** receiving HRSA grants.

American Society of Clinical Oncology (ASCO), Alexandria, VA, Jan-May 2017

ASCO's mission is to conquer cancer through research, education and promotion of the highest quality patient care. This data science project is an initiative of the ASCO's Informatics department which provides groundbreaking research and programs that makes a tangible difference in the lives of people with cancer.

Responsibilities:

- Data **massaging** and **mapping** -Using **MS SQL**, the Legacy Practice level information was migrated into current system field structures.
- **Normalizing** the data using SQL **pivot** tables and trending of normalized measures.
- **Dashboards** – Created visually impactful dashboards in **Power BI** and **Tableau** for data reporting. Extracted, **interpreted** and **analyzed** data to identify the **key metrics** and transform raw data into **meaningful**, **actionable** information.

Informatics & Data Analytics Projects Undertaken (2015-2017):

Predicting Hospitalization in Year -2 based on hospitalization in Year -1, Apr- May 2017

Environment used – Cloudera based Big Data ecosystem, using Hadoop, Spark and Spark SQL with HDFS and RDD concepts.

- Loading the claims data (claims, members, hosp2Y, labs) to HDFS and next loading the data for applying Data Mining Classification Algorithms using Spark to predict if patient is to be hospitalized in Year-2 from Year-1 history.

Predicting Drug Abuse in Emergency Admissions, Sep-Dec 2016

Environment used **–** Data used: **Substance Abuse and Mental Health Services Administration** (SAMHSA)-Drug Abuse Warning Network (DAWN) data from the 2005 through 2011. *R Programming Language, Oracle / SQL for descriptive analysis.*

- Using SQL and R, Multivariate regression models were constructed on Drug Abuse Warning Network (DAWN) data to predict drug misuse or abuse.
- Also, did a proof of concept in using SAS statistical tool.

Mining Associations between Drug abuse, Age and Suicide Attempts, Sep-Dec 2016

Environment used **–** Data used: Drug Abuse Warning Network (DAWN) data from the 2005 through 2011.Weka -Data Mining Software*, SQL for data preprocessing*

- Using Weka, Association mining (Market -Basket Analysis) algorithms Apriori and FP-Growth, were applied to the binarized data (using SQL), to find the most common combination of drugs used based on age category.

Smart Thermometer, Jan-May 2016

Environment used - Raspberry pi, Temperature sensor, Linux (OS), Python (Programming Language), Influx DB (Time series Database).

- Developed a prototype of a smart thermometer. Step1: An Infrared Sensor (IR) sensor measures the body temperature and relays it to the raspberry pi. Step2. Raspberry pi code interprets and analyzes the readings. Step3. Bluetooth transmits the information from pi to the phone app. Step4. The alarm in phone rings and alerts the parent/caregiver as per settings (high temperature).

Program on Machine Learning Algorithm for Breast Cancer Classification, Sept-Dec 2015

Environment used - Python (Programming Language), including-Pandas (Python Data Analysis Library), NumPy (scientific computing)

- Developed an algorithm using python (Pandas, NumPy, Matpotlib) based on a simple rule-based 'classifier' that can be used to predict the class (malignant or benign) of a set of unknown records, from the derived mean for each of 10 attributes describing the tumor.

Reduction of Hospital Readmissions, **Nov -Dec 2015**

Environment used – Tableau, Data Analysis and Visualization Software.

- Data analysis using Tableau, to identify constraints, aimed at reducing cost of hospital readmissions, thereby complying with Hospital Readmissions Reductions Program (HRRP) by the Centers for Medicare and Medicaid Services (CMS).

Wipro GE Healthcare, Chennai, India (2007-2011)

- Performed Data **Collection**, **Analysis** and Data **validation** by writing complex SQL **queries** using **SQL** server.
- Provided **analytic support** to patient quality improvement teams working in different Clinical Designs and formats data **reports** (simple to complex) utilizing SQL.
- Created functional **dashboards** from multiple data outlets.

KSR Institute of Dental Science & Research, India – *Research Associate, Technology Department (2005-2007)*

- Produced data and analysis for multiple departments in college to assist in decision making.
- Developed or enhanced several strategic models to assist in long term planning of health care.

Commendation:

Participant (Group of four) on behalf of George Mason University for the "Caring for the Caregiver Hack" Contest –March 2016.Conducted by: The Lindsay Institute for Innovations.

- Presented a concept about developing an application to help the Caretaker in taking care of the elderly.
- Commendation from sponsor of the event "Virginia Navigator"

Poster, on Mining Associations between Drug abuse, Age and Suicide Attempts

- *Environment used* **–** Data used: **SAMHSA data-Drug Abuse Warning Network (DAWN)** data from 2005 through 2011.Weka -Data Mining Software*, SQL for data preprocessing.*
- Presented the poster for The George Mason College of Health and Human Services, Celebration of Scholarship- May 2017.
- **Poster selected** by The American Medical Informatics Association, to submit the poster for **AMIA 2017 Annual Symposium, Washington D.C conference on Nov 6, 2017**.

Educational Qualification:

Master of Science in *Health Informatics and Data Analytics, class of May 2017, George Mason University, Fairfax, VA, G.P.A:3.78*

Bachelor of Dental Surgery*, Class of July 2005, Dr. M.G.R Medical University, Chennai, India, G.P.A:3.46*