KYI WIN

kyizarwin@outlook.com

(571) 268 -7889

www.linkedin.com/in/kyiwin

Address:

Vienna VA, 22182

EDUCATION

George Washington University, Washington, DC **May 2019**
M.S in Data Science

University of Virginia, Charlottesville, VA **Aug 2016**
B.A in Psychology

TECHNICAL SKILLS

Languages: R, Python (Pandas, Numpy, SciPy, Scikit-learn, Matplotlib, Seaborn), HTML, JavaScript, CSS
Operating Systems: Window, MacOS, Linux
Database System: MySQLWorkbench, Talend, MS SQL Server, Alteryx
Software: Tableau, SPSS, Anaconda(jupyter), Ataccama, Timeline, Plotly, Google Visualization API, datamatic, IBM
  Watson (Virtual Assistant), Microsoft Office Suite
Cloud: AWS (EC2, EMR with Hue, Hive and Pig), ELK stack, GCP
Statistic Knowledge: ANOVA, Linear Regression, Logistic Regression, Lasso Regression, Ridge Regression, K Nearest
  Neighbor, Decision Tree, Random Forest, K-mean, Perceptron, Support Vector Machine, Cosine Similarity, Multilayer
  Perceptron, Convolutional Neural Network

WORK EXPERIENCES

Accenture, Arlington, VA **Dec 2019- Feb 2020**
System Development Analyst
- Versed AWS with AWS Solution Architect Associate through the company learning portal-Cloud Guru and sandbox.
- Acquired many aspects of the account including previous projects the team accomplished and the current and recent projects the team is working on.
- Obtained Agile Certificate through Accenture analyst school.
- Participated in preparation of requesting the approval for the use of specific data tools from the client for an upcoming project.

Accomplishment: Created a mock up that would assist in succeeding the approval of a new data related tool.

Northrop Grumman, Mclean VA **May 2018- Aug 2018**
Data Science Intern*,* Enterprise Analytics
- Created a database model (Entity Relationship Diagram) to assist with obtaining required data from the database server.
- Developed several data analytic and visualization dashboards for complex large datasets.
- Collaborated with engineers, data scientists and enterprise managers for Enterprise Analytics project.
- Provided analytics and insights to stakeholders for decision making.
- Published a blog post on leveraging the visualization capabilities inside the Northrop Grumman data science community.

Accomplishment: Delivered data analytics and visualization dashboards to both technical and non-technical stakeholders.

University of Virginia, Charlottesville, VA **Aug 2015- July 2016**
Undergraduate Research Assistant, Initial Learnability and Familiarity with Icon Sequencing (AAC)
- Conducted a pilot study, which used both qualitative method- interviewing and collecting the participants' opinions, and quantitative method- recording the time it took for participants to finish a given test in the two different approaches: motor and semantic learning on 15 students.
- Involved in designing the experiment from planning stages to writing down detailed scripts to minimize experimental errors and obtained IRB-SBS Citi program training.
- Participated in weekly discussion of data analysis and discussed the ideas for further studies on the topic based on the results of the experiments.
- Analyzed the experiment's results using Excel and involved in preparation for poster presentation

Accomplishments: Involved in preparation of the research poster at the 9th Clinical AAC Conference in Curry School of
  Education, University of Virginia. https://pages.shanti.virginia.edu/Loncke_lab_Curry/minspeak/

DATA SCIENCE PROJECTS

Link: https://github.com/kyiwin

Cloud Computing
Lab1 - Launched an Amazon EC2 instance and set up a web-server on the instance using Flask, a lightweight web application. Created Load Balancing and Autoscaling groups.
Lab2 - Set up a Hadoop Cluster, access the Hue web interface, run MapReduce, Pig (locally and web interface) and Hive.
Lab3 - Established and applied Elasticsearch, Logstash, and Kibana for the Cyber Analytic.

Intro to Data Science
US 2012 Election - Determined the difference of the percentage of votes for the 2012 election. Applied ANOVA to determine if there is a difference in the percentage of votes for candidates in countries of small, medium, and large population using R.
Predicting Dementia - Modeled Logistic and Stepwise regression using R to determine the variables that have influence on predicting whether a person is demented or not.

Data Mining
Military Spending - Datamined to compare the military spending of top 20 GDP countries, and their growth rates.
Oil Price, Production, and the economy - Datamined to see if there are relationships between oil price, production and the economy.

Natural Language Processing
Mainstreet Bot- Applied NLP techniques to Mainstreet bank data and created sample chatbot using IBM Watson and slack bot application for consulting approach to customer experience with the website.

Machine Learning and Deep Learning
Income prediction - Used machine learning models to determine the important variables predicting income using US census data.
MINIST data - Gained hand-on experience on handwritten digit classification using multi-layer perceptron, and Street2shop data - classified on fashion items using convolutional neural networks.

Capstone Project
Novel Text Analysis using NLP techniques – Utilized R and R Shiny tools and NLP techniques to perform text analysis on the six books of the Lord of the Rings novel in search for how the novel is delivered to the audiences such as the way the author expressed and portrayed the characters, tone, style and the choice of words in artistic ways that it has reached up to 100 million copies of books being sold, and impact to them.

Data Warehousing
Trending on Social Media- Created database using Twitter and Youtube trending datasets collected using Social Feed Manager and Youtube API by utilizing Lucid chart and mySQL workbench in creating data architecture. Checked the data quality using Talent and analyzed the datasets using R and Python to find the trending.

Visualization of Complex Data
Education and Healthcare Spending- Exploration of Healthcare and Education spending using google visualization API combined with Html, CSS and JavaScript.
Flight Issue and popular airlines- Visualization of flight issues such as delay and safety, and determining popular airlines using revenue and passengers with Tableau.