

Kankshi Dhar

University of Maryland, College Park May 2020

✉ kankshi01@outlook.com

🌐 github.com/kankshidhar

☎ 857-241-7195

Education

University of Maryland - Master of Information Management

Data Analytics, Information Retrieval Systems, Data Science Models, Database Design

College Park, MD

Aug 2018-May 2020

University of Mumbai - Bachelor of Engineering in Computer Science

Relational Databases, Data Structures, Neural Networks, Artificial Intelligence

Mumbai, MH

Aug 2012-June 2016

Key Skills

- Programming Languages: Python, R, MySQL, Java, C++, Shell Scripting, NoSQL
- Software: Nltk, Tensorflow, Keras, Sklearn, Pandas, tidyverse, Statsmodels, Caret, dplyr
- Big Data: Hadoop, MapReduce, Spark, Hive, PostgreSQL
- Tools: Tableau, Power BI, ArcGIS, Git, Jira, Confluence, AWS Redshift, Sharepoint
- Analytics: SAS Enterprise Guide, Oracle Hyperion BI, Oracle DB, Clear Analytics, SAP BI

Work Experience

University of Maryland, Data Analyst

Aug 2018 - May 2020

- Enhancing data processing by creating complex queries to generate analytical insights for stakeholders & reducing run time by 26%
- Built ad-hoc reports & multifaceted requirement analysis for business teams leveraging BI tools such as SAS EG, SPSS
- Optimized hyper parameters with historical data for University metrics, boosting trend analysis using Tableau & Statistical Models
- Spearheaded FY19 and FY20 Financial impact profile creation for University aid & FAFSA trend analysis using Oracle BI
- Work with stakeholders throughout the organization to identify opportunities for leveraging company data to drive business solutions

Cheniere Energy Inc., Information Analyst Intern

May 2019 - Aug 2019

- Automated report generation while creating a unique data source to improve existing programs, boosting run-time by 6 hours
- Creating & Maintaining risk assessments for Information compliance programs w.r.t PHMSA & FERC governance policies for FY20
- Established a reusable trend tracker by defining key metrics to track business performance and Budgeting trends using historical data to project costs for FY 2020 using statistical Modeling and Excel tools such as Macros & Pivot Charts
- Monitor the compliance of processes and associated records to company standard operating procedure requirements

Reliance Jio Infocomm Ltd., SAP CRM Developer

Sep 2016 - Jun 2018

- Lead to the organization ranking #1 across the country by enhancing Enterprise services in a cross functional team to launch Individual owned products using ERP Suites such as SAP CRM & NoSQL databases - SAP HANA
- Designed & Managed databases with a focus on performance, security, scalability and disaster recovery for large Enterprise Customers
- Debugged existing ETL process and performed performance tuning on existing partner codes on AWS RedShift, reducing bugs by 18%
- Executed & evaluated user test cases for gap assessments, feasibility studies to assess cost/benefit and technical viability of DB solutions
- Increased Efficiency by designing Database views across multi-sourced customer information analysis during ERP product development

Research Projects

Text Mining & Classification for HireSmith Career Portal

Jan 2020 – May 2020

- Performing text mining & classification on Job descriptions for skill trend analysis in multiple industries within past 10 years
- Pre-processed data & applied TF-IDF with Topic modelling to analyze full text of job description and functions
- Identified 16 Industries & examined it's independent effect with 7 other categorised attributes across 5 million records
- Implemented Animated Tableau dashboards and plotly to display categorised attribute's trends across years

Medicare Rate Forecasting Model

Aug 2019 - Dec 2019

- Achieved 12% accuracy over previous models by developing logistic regression & Neural network model with ensemble methods to detect correlation, patterns & predict average Medicare provider's reimbursement rate in a dataset of 10 million records
- Conducted data pre-processing & augmentation to eliminate outliers, avoid overfitting & data exploration using Plotly, sklearn
- Models verified that highest paid provider will be plastic surgeons, Cardio-thoracic surgeons and Anesthesiologists in FY20

Football Database Design & Management

Jan 2019 – Apr 2019

- Developed a consolidated database to store, manipulate & search player & league statistics resulting in an efficient normalised system
- Designed an ER Model & translated model to relational schema using MySQL Database while implementing metadata management
- Created complex queries to generate dashboard views for leagues & match standings while performing diagnostic evaluations, best practice reviews and troubleshooting for an optimal Management system

Awards & Achievements

- Publication : "Offline Signature Corroboration Using Gray Level Features". International Journal of Engineering Science. Issue 3421
- Awarded with Best Sustainability Insight Award at the "Data Challenge 2020" for Ocean Cleanup Dataset by Booz Allen Hamilton
- Elected Director of Public Relations at Student Committee, 'YUVA' at University of Maryland (FY2019)