Sofia Dutta

Software developer, data analyst & data engineer

Sofiad1@umbc.edu • https://sofiadutta.github.io/

Summary

• 7 years of experience as a software developer, building systems for clients from all over the world

• big data analysis, deep learning, machine learning

Education

2019 – present master’s in data science, umbc, baltimore, usa 4.0/4.0

2006 – 2010 bachelor’s in computer science and engineering, wbut, india 3.5/4.0

Data science projects

• comparison of word2vec and doc2vec model driven sentiment analysis using svm, lr, keras cnn, bidirectional lstm with and without pre-trained word and document embeddings.

• amazon dataset: https://sofiadutta.github.io/datascience-ipynbs/sentiment-analysis/sentiment\_analysis\_amazon\_laptop\_review.html

• imdb dataset: https://sofiadutta.github.io/datascience-ipynbs/sentiment-analysis/sentiment\_analysis\_imdb\_movie\_review.html

• data characterization of baltimore city employee salaries to study income inequality in government link: https://sofiadutta.github.io/datascience-ipynbs/eda/data\_analysis\_baltimore\_city\_salaries.html

• data characterization of nyc film permits to figure out where my favorite films being shot at

Link: https://sofiadutta.github.io/datascience-ipynbs/eda/data\_analysis\_nyc\_film\_permits.html

• data characterization of nyc fire department data to find impactful events via a data analytics path

Link: https://sofiadutta.github.io/datascience-ipynbs/eda/data\_analysis\_nyc\_fire\_department.html

• big-data analytics and sentiment classification with twitter stream data. See project links below:

• https://sofiadutta.github.io/datascience-ipynbs/big-data-analytics/using\_myclassifier\_twitter\_data\_sentiment\_classification\_and\_big\_data\_analytics\_on\_spark\_dataframe.html

• https://sofiadutta.github.io/datascience-ipynbs/big-data-analytics/twitter\_stream\_analysis\_using\_spark\_streaming.html

• https://sofiadutta.github.io/datascience-ipynbs/big-data-analytics/using\_sparkmllib\_word2vec\_model\_driven\_logreg\_sentiment\_myclassifier\_creation.html

Technical skills

Languages python, java, sql, pl/sql

Tools & stacks scikit-learn, mllib, pyspark, pl/sql developer, oracle report/forms/workflow builder

Back-end oracle (9i, 10g, 11g), mongodb, json

Enterprise tools google cloud, aws s3, oracle fusion 11, oracle applications (r12 & 11i)

Work experience

2019 - present student researcher in semantic web, ebiquity research group, umbc, usa

2018 – 2010 software consultant and data analyst at tata consultancy services (tcs) ltd., india

Tcs projects

Jan’17 - feb’18 commercial bank of dubai, uae

• was a key project resource trained to work with oracle fast formulas

• led a team of developers working on oracle fusion human capital management

• prepared requirement analysis, software design and test plans documents

Sep’16 - nov’16 kaiser permanente, usa

• led a team of developers working on oracle fusion (hcm) functionalities

• prepared requirement analysis, software design and test plans documents

Feb’16 - jun’16 hyatt-oracle r12 rollout, usa

• ensured efficient execution of data queries using pl/sql stored procedures.

• worked on ach payment, invoice management and hrms configuration.

• performed requirement and change analysis.

Jun’15 - feb’16 staples, usa

• ensured efficient execution of data queries using pl/sql stored procedures.

• worked on ach payment, invoice management and hrms configuration.

• prepared requirement analysis, software design and test plans documents

Jan’14 - jun’15 dhl, uk

• ensured efficient execution of data queries using pl/sql stored procedures.

• handled hrms and payroll configuration and administration.

• seamlessly managed data migration activities from oracle apps to erp suite.

• tested system functionality and bug fixes.

Sep’13 - jan’14 tata steel, india

• ensured efficient execution of data queries using pl/sql stored procedures.

• prepared requirement analysis, software design and test plans documents

May’13 - sep’13 ihcl - r12 reimplementation, taj hotel, mumbai

• ensured efficient execution of data queries using pl/sql stored procedures.

• handled change management activities for oracle reports, forms and databases.

• seamlessly managed data migration activities.

• prepared requirement analysis, software design and test plans documents

Apr’13 - may’13 cmcl - oracle erp r12 implementation, star cement, kolkata

• ensured efficient execution of data queries using pl/sql stored procedures.

Jun’11 - apr’13 bangladesh railway oracle apps, bangladesh railway, bangladesh

• ensured efficient execution of data queries using pl/sql stored procedures.

• developed report generation packages for payroll management

• prepared requirement analysis, software design and test plans documents

Certifications completed

• oracle 9i sql (1z0-007)

• oracle 9i pl/sql (1z0-047)

• oracle database 11g: advanced pl/sql (1z0-146)

• oracle e-business suite 12 financial management implementation specialist: receivable (1z0-518)

• oracle fusion hcm base cloud service 2016 implementation essentials (1z0-329)

• oracle global human resources cloud 2017 implementation essentials (1z0-965)

• oracle talent management cloud 2017 implementation essentials (1z0-966)