

## SOFIA DUTTA

Aspiring Data Scientist and Machine Learning Engineer | Email: [sofiad1@umbc.edu](mailto:sofiad1@umbc.edu)

### SUMMARY

- Aspiring Data Scientist and Machine Learning Engineer
- 7 years of experience in Software Consultancy and Data Analysis

### EDUCATION

Current – 2019	Master's in Data Science, UMBC, Baltimore, USA	4.0/4.0
2010 – 2006	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.
  - 📖 Sentiment Analysis on Amazon laptop reviews: [https://sofiadutta.github.io/datascience-ipynbs/sentiment-analysis/Sentiment\\_Analysis\\_Amazon\\_Laptop\\_Review.html](https://sofiadutta.github.io/datascience-ipynbs/sentiment-analysis/Sentiment_Analysis_Amazon_Laptop_Review.html)
  - 📖 Sentiment Analysis on IMDb movie reviews: [https://sofiadutta.github.io/datascience-ipynbs/sentiment-analysis/Sentiment\\_Analysis\\_IMDB\\_Movie\\_Review.html](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](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](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](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/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/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](https://sofiadutta.github.io/datascience-ipynbs/big-data-analytics/Using_SparkMLLib_Word2Vec_Model_Driven_LogReg_Sentiment_MyClassifier_Creation.html)

### TECHNICAL SKILLS

Platforms	Mac, Unix, Windows
Languages	Python, Java, SQL, PL/SQL
Tools & stacks	Scikit-learn, MLlib, PySpark, PL/SQL Developer, Oracle Report/Forms/Workflow Builder
Back-end	MongoDB, Oracle (9i, 10g, 11g)
Enterprise tools	Google Cloud, AWS S3, Oracle Fusion 11, Oracle Applications (R12 & 11i)

### WORK EXPERIENCE

Current – 2019	Student researcher in Semantic Web and Smart Home Access Control, Ebiquity Group, UMBC, USA
2018 – 2010	Software Consultant and Data Analyst at Tata Consultancy Services (TCS) Ltd., India

### TCS PROJECTS

Feb'18–Jan'17	Commercial Bank of Dubai, UAE <ul style="list-style-type: none"> <li>• Was a key project resource trained to work with Oracle Fast Formulas</li> <li>• Led a team of developers working on Oracle Fusion Human Capital Management (HCM) functionalities like resignation, transfer, promotion, performance rating etc.</li> <li>• Prepared software design and test plans documents</li> </ul>
---------------	--

	<ul style="list-style-type: none"> <li>Performed requirement and change analysis.</li> </ul>
Nov'16–Sep'16	<p>Kaiser Permanente, USA</p> <ul style="list-style-type: none"> <li>Led a team of developers working on Oracle Fusion (HCM) functionalities</li> <li>Prepared software design and test plans documents</li> <li>Performed requirement and change analysis.</li> </ul>
Jun'16–Feb'16	<p>Hyatt-Oracle R12 Rollout, USA</p> <ul style="list-style-type: none"> <li>Ensured efficient execution of data queries using PL/SQL stored procedures.</li> <li>Worked on ACH payment, invoice management and HRMS configuration.</li> <li>Performed requirement and change analysis.</li> </ul>
Feb'16–Jun'15	<p>Staples, USA</p> <ul style="list-style-type: none"> <li>Ensured efficient execution of data queries using PL/SQL stored procedures.</li> <li>Worked on ACH payment, invoice management and HRMS configuration.</li> <li>Performed requirement and change analysis.</li> </ul>
Jun'15–Jan'14	<p>DHL, UK</p> <ul style="list-style-type: none"> <li>Ensured efficient execution of data queries using PL/SQL stored procedures.</li> <li>Handled HRMS and Payroll configuration and administration.</li> <li>Seamlessly managed data migration activities from Oracle Apps to ERP Suite.</li> <li>Performed requirement and change analysis.</li> <li>Tested system functionality and bug fixes.</li> </ul>
Jan'14–Sep'13	<p>Tata Steel, India</p> <ul style="list-style-type: none"> <li>Ensured efficient execution of data queries using PL/SQL stored procedures.</li> <li>Worked on software design and test plans documents</li> </ul>
Sep'13–May'13	<p>IHCL - R12 Reimplementation, Taj Hotel, Mumbai</p> <ul style="list-style-type: none"> <li>Ensured efficient execution of data queries using PL/SQL stored procedures.</li> <li>Handled change management activities for Oracle reports, forms and databases.</li> <li>Seamlessly managed data migration activities.</li> <li>Worked on software design and test plans documents</li> </ul>
May'13–Apr'13	<p>CMCL - Oracle ERP R12 Implementation, Star Cement, Kolkata</p> <ul style="list-style-type: none"> <li>Ensured efficient execution of data queries using PL/SQL stored procedures.</li> </ul>
Apr'13–Jun'11	<p>Bangladesh Railway Oracle Apps, Bangladesh Railway, Bangladesh</p> <ul style="list-style-type: none"> <li>Ensured efficient execution of data queries using PL/SQL stored procedures.</li> <li>Developed report generation packages for payroll management</li> <li>Worked on software design and test plans documents</li> </ul>

## CERTIFICATIONS COMPLETED

- o Oracle 9i SQL (1Z0-007)
- o Oracle 9i PL/SQL (1Z0-047)
- o Oracle Database 11g: Advanced PL/SQL (1Z0-146)
- o Oracle E-Business Suite 12 Financial Management Implementation Specialist: Receivable (1Z0-518)
- o Oracle Fusion HCM Base Cloud Service 2016 Implementation Essentials (1Z0-329)
- o Oracle Global Human Resources Cloud 2017 Implementation Essentials (1Z0-965)
- o Oracle Talent Management Cloud 2017 Implementation Essentials (1Z0-966)