

## **Snehalata Deorukhkar**

**E-mail:** snehalata.deorukhkar@gmail.com

**Mobile:** +91 9820825922

### **Objective:**

---

To take up challenging assignments by which my knowledge, skill and experience are leveraged to the full extent and contribute to the growth and development of the organization through creative and innovative ideas.

### **Summary:**

---

- ~11 years of experience in IT with strong skills in Hadoop, Machine learning and Statistics and ~5 years of experience in Data Science
- Analytical and detail-oriented with excellent problem-solving skills
- Result oriented with excellent communication & people skills.
- Oracle Certified Java Programmer.
- Hortonworks certified Apache Hadoop developer.
- Was speaker at Hadoop Summit, 2014 held at San Jose, California. The topic of the presentation was 'Big Data Analytics – Open Source Toolkits'

### **Career Chronology:**

---

<b>Designation</b>	<b>Company</b>	<b>Duration</b>
Data Scientist	General Mills	Feb-2016 – Till Date
Project Lead	Syntel Ltd.	May-2015 – Feb-2016
Sr. Software Engineer	Syntel Ltd.	Apr-2013 to Apr-2015
Software Engineer	Syntel Ltd.	Apr-2011 to Apr-2013
Graduate Engineer Trainee	Syntel Ltd.	Apr-2010 to Apr-2011
Junior Lecturer	K.J.Somaiya College Of Engineering, Vidyavihar, Mumbai	July-2009 to Apr-2010

### **Technical Skills:**

---

Reporting Tools	iReport, Jasper Server, Cognos 8, PowerBI
DBMS/RDBMS	MS SQL Server, Oracle SQL Developer, MySQL
Languages	Python, Core Java, JDBC, SQL, PL/SQL, GWT
Big Data Technologies	Hadoop, MapReduce, Hive, Pig, MongoDB, Spark
Machine learning	Linear Regression, Classification, Segmentation, Topic Modelling, Sentiment Analysis, Time Series Forecasting

## Academic Qualification:

---

- B. E. (Computer) from Fr. Conceicao Rodrigues College of Engineering, Bandra, Mumbai with 61% in final year
- H.S.C from D.G. Ruparel College, Mahim, Mumbai with 81.17%
- S.S.C. from Vidyamandir Highschool, Vikhroli with 90.93%

## Details Of Project Worked:

---

**Company Name:** General Mills

**Role:** Data Scientist

**Responsibilities:**

Developing and implementing analytical solutions, innovating and automating analytical outputs using machine learning and mentoring teammates.

**Major projects:**

1. **Consumer Lifecycle:** This project involved researching lifetime of consumers visiting our recipe sites. Presented findings and gave suggestions on areas to focus on to increase the consumer retention.  
Tools and Technologies used: Python, Scala, Spark, Hive, Excel
2. **School menu analysis:** CnF K-12 team wanted to understand what schools are putting on school menu to ensure we are delivering the right meal solutions to schools, identify differences by school grades and find opportunities for growth. Started with no data and scraped menu aggregators sites to create data for analysis. Created menu hierarchy using multi-label classification model. Matched school names to school population in US census data using tf-idf and cosine similarity. Self-serve dashboard was created using Tableau.  
Tools used: Python, Selenium, Excel, Impala, Tableau
3. **Category Forecasting:** Worked on forecasting category growth using multiple regression model using Nielsen point of sales data.  
Tools used: R, Excel
4. **Yogurt segmentation:** Worked on demand space segmentation on the usage and attitude study survey data for yogurt.
5. Also worked on text analysis using topic modelling and sentiment analysis, automated social data pulling using crimson hexagon api and ad-hoc statistical and exploratory data analysis

**Duration:** From Feb, 2016 – Till Date

---

**Company Name:** Syntel

**Role:** Technical Lead - Hadoop

**Responsibilities:** Assigning tasks among team members and tracking progress on each of the tasks assigned. Helping team members in resolving issues faced in their respective tasks. Designing the approach for migrating Netezza scripts in pig and hive scripts. Training team members on Hadoop enabling them to work independently in Hadoop.

## Major projects:

1. **GMM (Global Marketing Measurement)** is Reporting platform for AMEX Offers to measure performance of Digital campaigns to improve campaign marketing. Migrated the existing GMM application designed using Netezza and DataStage into Hadoop. Did Performance optimization of pig scripts to gain performance benefits over Netezza scripts.  
Tools used: Hadoop, Pig, Hive, UNIX Shell Scripting, Kognitio
2. **EDW-Hadoop Optimization PoC:** The long running jobs of the Oracle EDW were identified. A load script from one of the long running jobs was migrated into Hadoop-Hive to demonstrate performance benefits with Hadoop. Created source tables in Hive and loaded data from Oracle tables to Hive using Sqoop. Migrating Oracle-SQL script into HiveQL and performance optimization of Hive script  
Tools used: Hadoop, Hive, Sqoop
3. **Big data practice:** Worked on various PoCs like usage-based insurance, price optimization, Customer360. Conducted multiple Hadoop trainings for SOG academy. The training included theory lectures and practical sessions covering HDFS, MapReduce, Hive and Pig.  
Tools used: Hadoop, Hive, Sqoop, PHP
4. **Actuate to Jasper Migration:** The project involved migrating State Street's existing reporting system, which was implemented using Actuate reporting tool, Actuate iServer and Actuate APIs, using jasper reports and java.  
Tools used : iReport, Jasper Server, Core Java

**Duration:** From April, 2010 – Feb, 2016

---

## Personal Information:

---

Date of Birth	09-Nov-1986
Gender	Female
Marital Status	Married
Passport No.	H1003075
Visa	B1 (Valid Till 6 <sup>th</sup> May 2024)
Languages Known	English, Marathi, Hindi
Address	E-402, Asha Milap, Samarth Garden CHS, Bhandup West, Mumbai – 400 078