# ROSY GUPTA

**Address for Correspondence**
University of Petroleum and Energy Studies
Energy Acres PO Bidholi via Prem Nagar
Dehradun-248007 (Uttarakhand) – India
**Phone: +91-9557629654**
**Email:** rosy01.gupta@gmail.com          **Linkedin**: https://in.linkedin.com/in/rosygupta

## CAREER OBJECTIVE

To embark on an educational journey in the research areas of Big Data Analytics, Data Mining, Machine Learning and Distributed Computing that offer both a challenge and an opportunity for growth tantamount to my ability and technical knowledge and to expand my skills and interest through global research exposure.

## ACADEMIC BACKGROUND

| Year(s) | Qualification-Degree/ Diploma/ Certificate | Board/ University | College/ Institute | Percentage/ CGPA |
|---|---|---|---|---|
| 2012-16 | Bachelor of Technology in Computer Science (specialization in Oil and Gas Informatics) | University of Petroleum and Energy Studies, Dehradun | College of Engineering Studies | 3.40/4 |
| 2010-2011 | Diploma in Management-Level 5 | New Zealand Institute of Management | Kiwi Institute of Training and Education (KITE), Auckland | 90% |

## TRAININGS AND PROJECTS UNDERTAKEN

**Major Project**                                                        **Duration: August, 2015 – December, 2015**
**Project Title:**   Hack the Talk
**Mentor:**   Dr. Kingshuk Srivastava, Asst. Professor, UPES
**Description:**   Implemented classification of emotions based on spectral features using the **k-nearest neighbour** and **Support Vector Machines** (SVM) as classifiers. The performance of the classifier was evaluated for **real life speech corpuses** (telephonic conversations) in **multiple languages** – Hindi, Tamil, Telegu and English.

**Institution:**   **University of New Brunswick, Canada**          **Duration: 12 weeks, May, 2015-August-2015**
**Project Title:**   **Real Time Energy Data Collection and Mining using OpenTSDB –SG2B Electric**
**Mentor:**   **Dr. William McIver, Jr., NSERC Industrial Research Chair**
**Description:**   The aim of the industry project is to develop a database structure capable of collecting **real-time minute-to-minute time-series data** from millions of housing energy consumption and provide advice to SG2B on using this new type of database in their day-to-day operation. The **Smarthomes API** was used to feed the real-time time series data to **OpenTSDB**, **HBase** and **HDFS**. A feeder demand forecasting function was developed using historical data and weather information. Displayed the results on website using **Javascript Highchart**.

**Institution:**   **University of New Brunswick, Canada**          **Duration: 12 weeks, May, 2015-August-2015**
**Project Title:**   **Mobile First Technology Initiative**
**Mentor:**   **Dr. William McIver, Jr., NSERC Industrial Research Chair**
**Description:**   The aim of the project is to create anytime, anywhere user **mobile experience**. Three topics of acute interest within mobile computing sector were identified: interoperability, improving consumer experience and assessing and improving security models that are used on mobile devices.

| | |
|---|---|
| **Minor Project-II** | **Duration: January, 2015-May-2015** |
| **Project Title:** | **Analytics as a Service (AaaS) on OpenStack** |
| **Mentor:** | **Dr. Kingshuk Srivastava, Asst. Professor, UPES** |
| **Description:** | Integration of **OpenStack** with **Hadoop** through data processing and thereby managing Hadoop clusters within the cloud. Provision and manage Hadoop, **Spark** and **Storm** clusters on Openstack. Thereby, faster provisioning of data processing clusters on Openstack. |

| | |
|---|---|
| **Minor Project-I** | **Duration: August, 2014 – December, 2014** |
| **Project Title:** | **Crude Oil Economic Evaluation and Prediction Modelling** |
| **Mentor:** | **Dr. Kingshuk Srivastava, Asst. Professor, UPES** |
| **Description:** | Analyze the existing studies of prediction models for crude oil prices and thereby develop a software system for real-time crude price analysis and prediction. Two different algorithms for attribute selection i.e. **Sequential Minimal Optimization** for regression and **Instance-Based Learning** method were used. The data was sent to Hadoop Distributed File System (**HDFS**) and processed using **MapReduce** programs in Python. The data visualization was also done using **Python Graphs**. |

**Other Projects**

| | |
|---|---|
| **Project Title:** | **Twitter Sentiment Analysis- FIFA World Cup, 2014** |
| **Description:** | Implementation of **Hadoop multinode cluster**. Thereby, conducted real-time Twitter sentiment analysis for the FIFA World Cup. The system was designed to gather data from the **Twitter Streaming API**, send the tweets to files on HDFS through **Flume**, and use **Oozie** to periodically load the files into Hive, where we could query the raw **JSON data**, through the use of a **Hive SerDe**. |

## TECHNICAL CERTIFICATIONS

- Completed **Scalable Machine Learning** course at EDX by University of California, Berkley (2015).
- Completed **Introduction to Big Data and Apache Spark** course at EDX by University of California, Berkley (2015).
- Certified **EMC Academic Associate** in Information Storage and Management (2015)
- Certified **XEBIA Hadoop Developer** (2015).
- Course completed for **Python Training** at Spoken Tutorials, IIT Bombay (2014).
- Completed **Data Scientist's Toolbox** course at Coursera by John Hopkins University (2014).

## ACTIVITIES AND AWARDS

- Received **MITACS Globalink Research Scholarship Award 2015** for summer research internship in Canada.
- Zonal level winner of **India's Biggest Networking Championship (IBNC)** (2014) and hence qualified for the National Level Workshop of Cloud Computing held at IIT Delhi in March 2015.
- Best Student Award for Diploma in Management at KITE, Auckland, New Zealand.(2014)
- Committee Member of UPES **Google Hub** (2014-15)
- Served as a Committee Member in **Public Relations Committee** for UPES CSI Student Chapter. (2013)
- Won **Ms. Techie** title at Mr. and Ms. Techie organised by UPES CSI Student Chapter in 2012.

## COMPUTER SKILLS

- **Programming Languages:**   Python**,** C, C++, Java, Shell Scripting
- **Big Data Technologies:**   Apache Hadoop, Hive, Pig, Apache Flume, Oozie, HBase, OpenTSDB
- **Databases**:   MySQL, IBM DB2, Mongo DB, Oracle (PL/SQL)
- **Web Technologies:**   HTML5, CSS3, Javascript, SQL, PHP, AJAX, XML
- **Operating Systems**:   Windows, Linux (Ubuntu 14.4)
- **Others:**   Git, Apache2

## PERSONAL DETAILS

| | | |
|---|---|---|
| Gender | : | Female |
| Nationality | : | Indian |
| Permanent Address | : | House No. 16, Sector 18, Panchkula, Haryana, India-134109 |
| Languages Known | : | English, Hindi and Punjabi |