# Rubi Pandey

rubypandey92@gmail.co (M) +91 8054396300
1694, Sec-29 B, Chandigarh fb.com/ruby.pandey.98

***Senior Analyst in Data Science at Eclerx Services Ltd with experience in Machine Learning and NLP***

## Basic Info

30th Jun 1991
Chandigarh
rubypandey92@gmail.com

## Personal Skills

Analytical Skills

Communication Skills

Time Management Skills

## Interests

- Painting
- Travel
- Tennis
- Camping
- Swiming

## Objective

To excel in the field of data science and build applications using AI and machine learning and contribute towards the growth of my organization.

## Summary of Experience

| | |
|---|---|
| 2013-14 | **Research Intern (Data Science) at Technomate Solutions**<br>My role was to complete various projects related to Python and Machine Learning |
| 2014-15 | **Freelancer**<br>I used to work for Freelance Projects related to machine learning, Text Mining and Predictive Analytics |
| 2017-present | **Senior Analyst (Data Science) at Eclerx Services Ltd.**<br>Working on projects related to Data Science and NLP for BFSI Clients. |

## Education

Masters in Engineering in Power Systems (2015-2017)
**University Institute of Engg & Technology, Chandigarh**
Dissertation in Machine Learning applications on Smart Grid Data

MBA in Finance & IT (2012-2013)
**University Institute of Engg & Technology, Chandigarh**
Project in Investment Banking

B.E. in Electrical & Electronics Engg (2008-2012)
**University Institute of Engg & Technology, Chandigarh**

## Technical Skills

### Machine Learning

Have understanding and experience in writing codes of various machine learning algorithms for classification, clustering, regression, Neural Networks, ANFIS, Feature Selection, Dimensionality Reduction using PCA

### Data Scrapping & Web Scrapping

Have worked with data scrapping tools like Tabula and BeautifulSoup. Have worked with extracting data from webpages and different format files.

### Natural Language Processing

Worked on projects related to text mining involving data pre-processing, TF-IDF approach, entropy reduction, sentiment analysis, content retrieval from electricity grid reports, etc.

### Optimization

Experience in working with mathematical optimization techniques like gradient descent, numerical techniques and meta-heuristic algorithms like PSO.

### Artificial Intelligence

Utilized and written codes for AI algorithms like Neural Network, RNN, and designed a hybrid technique using RNN, Text Mining and Machine Learning for Load Forecasting in Smart Grid.

### Information Extraction

Worked with development of solutions for Information Extraction from Financial Documents. Worked with several kinds of data like html, pdf, excel etc.

## Programming Languages

Efficient with Python and MATLAB. Have medium level expertise in C/C++

## Tools Used

| | |
|---|---|
| Modeling Tools: | Expertise with MATLAB, SIMULINK, ANFIS |
| NLP Tools: | Worked with NLTK, pandas, numpy, scipy |
| Machine Learning Tools: | Worked with Sci-kit Learn, BeautifulSoup, etc. |
| Optimization Tools: | Worked with GAMS and HOMER for Algaebraic Modeling |
| GUI Tools: | Form Designs using Visual Basic, GUI development using tkinter in Python |
| Database Tools: | Worked with MySQL and connectivity of SQL with Python using pymysql |

Master's Project

### Islanding detection in Smart grids using Machine Learning Algorithms.

The project was completed as my Master's Dissertation and solves the problem of Islanding detection which occurs in the presence of Distributed generators in Smart Grid. It utilized Ensemble learning methods to present a predictive model for Islanding detection.

### A Hybrid approach of AI, Machine Learning and Text Mining for Short Term Load Forecasting in Smart Grid.

The project was completed as my Master's Mini Dissertation and solves the problem of STFL in Smart Grid using a Recurrent Neural Network Approach. To improve the prediction accuracy several other factors (weather information and smart grid parameters) are obtained by mining Electricity consumption reports and Weather Websites and Feature Extraction is applied to the enhanced data which is then fed to the RNN. Using this hybrid approach enhanced the prediction accuracy from 79 to 84%.