

---

# Dattatraya Walunj

## Data Analyst(Data Scientist)

Shitole Nagar, Old Sangvi  
Pune, 411027  
9763529046  
drwalunj.2010@gmail.com

### SUMMARY

1 year of experience in Python(Django) and 6 months of experience in Data Analyst(Data Science). Quick learning abilities to adapt to current technology and market needs. Experience in machine learning problem solving during projects in Internships. Experience in data science libraries like NumPy and Pandas with thorough knowledge of python visualization in Seaborn, Plotly, Matplotlib, Keras, Tensorflow and Tableau.

### SKILLS

**Programming Languages:** Python, Beginner level C, C++, Java, PHP, SQL.

**Frontend Technologies:** HTML, CSS, JavaScript

**Backend Technologies:** Django, Flask (Python)

**Operating Systems:** Linux, Windows

**Databases:** MySQL, PostgreSQL, SQLite, MongoDB

**Others:** Git, Bitbucket, Sourcetree, Postman (API development), spaCy, Scikit-Learn, SciPy, Natural Language Processing (NLP) and Deep Learning.

### EXPERIENCE

#### **RapidAI Technologies, Pune** - *Data Analyst (Data Scientist)*

SEP 2019 – MAR 2020

- **Resume Parsers:** A simple resume parser used for extracting information from resumes.
- The main challenge is to extract names from resumes. I used a more sophisticated tool called spaCy. spaCy comes with Pre-trained models for tagging, parsing and entity recognition.

- 
- To extract phone and email from a resume used a regular expression model (re.findall). To extract skills from a resume used Sentence Tokenization and Word Tokenization Technique.
  - For extracting Education Details from resumes use the NLTK module.

### **Brainfuse, New York** - *Online Tutor/Expert*

SEP 2017 – PRESENT

- Assessing the difficulty that the student is having with the work.
- Designing/finding a way to guide the student and implementing it.
- Communicating throughout the session.

### **Techknit IT Enabled Services Pvt. Ltd, Pune** - *Online Tutor/Expert*

AUG 2013 – SEP 2017

- Clarifying the Student's Needs.
- Choosing an Effective Approach.
- Ensuring Student Understanding.
- The experience of understanding another person's condition from their perspective.

## **PROJECT**

### **Resume Formatting:**

Parsing experience related details from the resume and rewrite it with defined destination formatting.

- Parsed .doc, .docx, and .pdf file formats.
- Project done using python.
- Here the thing that is captured from each company given in the resume. I captured the Job title, Name of the company, City, State, Start date and End date.

### **Credit Risk Analysis** - *Loan Default Prediction*

OBJECTIVE:

- Create a predictive model to classify each borrower as a defaulter or not using the data collected when the loan has been given.
- Minimize the risk of borrowers defaulting on loans using a created model.

PROBLEM STATEMENT:

- To classify if the borrower will default the loan using borrower's finance history. That means, given a set of new predictor variables, we need to predict the target variable as **1 => Defaulter** and **0 => Non-Defaulter**.

---

#### EVALUATION APPROACH:

- I used a classification algorithm for solving our problem. It establishes a relation between input features and output features.
- For the Loan default prediction project, False Negatives Rate is the best metric to evaluate the model. Lower the number of false negatives, better the model is.
- From the confusion matrices, I infer that the Logistic Regression model had the least False Negative Rate (FNR).

## EDUCATION

**Department of Mathematics, University of Pune - *M.Tech***

JUNE 2010 - JULY 2013

**Modern College, Pune(Shivajinagar) - *B.Sc(Mathematics)***

JUNE 2009- JUNE 2010

## MY WEBSITES

<https://www.pdgmpx.com>

<https://airesumeparser.herokuapp.com>

<https://github.com/Dattatraya77/Corona-Patient-Prediction>