

# Pritam Bhattacharjee

**Data scientist with expertise in commercial NLP, NLU, NLG, computer vision and predictive analytics in python**

Howrah, West Bengal

[pritambhattacharjee56\\_c4u@indeedemail.com](mailto:pritambhattacharjee56_c4u@indeedemail.com)

+91 8697047041

Hi ,

I am Pritam, a 6 years of experienced expert data scientist of applied AI research with expertise in commercial NLP , computer vision and predictive analytics projects in python 3.6 and 3.7.

I have data science projects clients from USA, Canada, Japan and india with their happiness on my completed work.

AS A MENTION,

SOME OF MY IMPORTANT WORK EXPERIENCES ARE -

1.

I worked as co-founder and R&D CHIEF AI and top Manager on financial internet data parsing, NLP and prediction in an AI financial prediction solution startup company from USA of an angel investor , invested in multi companies. I leaded 10 interns from

""Harvard and MIT as masters and phd""

in this company and done multiple projects simultaneously for better performance and also done interviews to hire interns for this company. AND I LEARNED LOT OF MANAGEMENT SKILLS OF AI PRODUCT DEVELOPMENT FROM THE ANGEL INVESTOR AS A CO-FOUNDER.

2

I worked as lead AI Manager for interns for an big potential AI event marketplace startup from USA to build their recommendation system and lead their AI intensity team.

3

I worked as a R&D AI lead of NLP, NLU, NLG, computer vision and predictive analytics for a big potential online workpkace and venue location marketplace startup from UK.

4

I worked for a geospatial solution company for satellite prediction of soil vegetation.

5

I worked for a German company on search technology and docker system improvement of the companies core product and done several fixes and provided interesting ideas with implementation option.

6

I worked for a supply chain AI automation company from USA and selected to lead 4 interns in this company.

7

I worked as an CHIEF AI and Lead AWS cloud engineer for financial fraud detection AI company from India with very big customers from USA and Africa.

8

I worked for 1 year in a big financial project as a portfolio optimizer program with dataset pre-processing to be ready for optimization on close price column and calculated log returns and shrinkage return for long short PORTFOLIO optimization , pre-filter by evolutionary algorithm and after that

to select date range from start and end calendar box, input typed days returns and risk free rate, select candidates selection metrics from about 33 hybrid and single metrics including treynor, ULCER INDEX, sterling, burke, alpha, beta, sharpe, systematic risk, unsystematic risk and many other and MENTIONABLE latest implementation of rachev version 1 and rachev percentile version and farinelli tibletti aggressive and conservative metric versions with hybrid on these from RECENT DEVELOPMENT RESEARCH PAPERS and then single and about 6 hybrid balancing as targets for efficient frontier optimization with visualization and result table with weights and about 8 metrics and these products as windows 10 GUI desktop application for by tkinter, pandastable, tkcalendar, pyinstaller and lot of other python libraries as exe product for 1 of my client from USA .

9

I worked for 1 month in a covid 19 project and built a covid 19 x-ray image recognition ipynb file to py file to tkinter desktop GUI application using pyInstaller python library with lot of configuration to build exe compatible trained model file and the trained accuracy was 92 % and validation accuracy is 100 % with roc\_auc score more than 92 % and raw accuracy, specificity and sensitivity are each 100 % .

10

I was got invitation to build recommendation system by a company from Canada as co-founder position and by a company from India as CTO position.

My very important upgraded AI data science skills with implementation experience in python 3.6 and 3.7 are

1.

VERY DEEP DIVE EXPERIENCE AND UNDERSTANDING OF TRANSFORMERS AND APPLY IT FOR ANY TASK OF NLP, NLU, NLG AND COMPUTER VISION.

And

2

Image search engines, text search engine, custom or NEW object detection model train with pre-trained model fine tune and detect target objects by loading the fine tuned custom object trained model.

3

Build state-of-the-art visualizations using tableau dashboard.

4

Product setup and launch on AWS cloud and Google cloud. I already build end to end get and post API with authentication for a big potential AI startup from India as this company 's core infrastructure on AWS and tested thoroughly multiple time the total system from python code automation and postman and these all worked successfully. So, I have real implementation experience on AWS for end to end and other projects.

Also I built object detection models on Google cloud with lot of processing on annotation file to csv format and the trained model have above 99 % accuracy for 2 category objects detection.

AS A MENTION, I AM EXPERT AT ANY TASK OF NLP SUCH AS CHATBOT BUILD, QUESTIONS ANSWERS, TEXT TRANSLATION, ENTITY RECOGNITION, TOPIC MODELING, TEXT CLASSIFICATION, TEXT GENERATION, SENTIMENT ANALYSIS, TOPIC EXTRACTION, OCR ON IMAGE AND LOT MORE OF ANYTHING YOU CAN IMAGINE OF NLP, NLU AND NLG WITH BETTER ACCURACY ABOVE 90 % WITH TRANSFORMER THAT IS CORE ENGINE OF ALL STATE-OF-THE-ART NLP, NLU AND NLG MODELS IN GPT3, GPT2, BERT, ROBERTA, HUGGINGFACE, LONGFORMER, SPARK NLP AND MORE.

LONGFORMER IS A MODEL THAT CAN PROCESS INPUT OF 10000 WORDS AS IN THIS TIME OF EXTENSIVE USE OF BIG DATA ACROSS INDUSTRIES. BECAUSE OTHER MODELS ALTHOUGH MAY RECEIVE INPUT OF 10000 WORDS, BUT ONLY FIRST 768 TO 1000 WILL BE USED FOR PREDICTION AND

THIS MEANS CORRECT PREDICTION WILL NOT BE RECEIVED FOR DIFFERENT CONTENT AFTER 768 TO 1000 WORDS IN INPUT. AND I CAN FIX THESE USING THE LONGFORMER MODEL.

- 

Also I have some advanced innovative ideas on developing hugely massive search space data with very faster result analysis and retrieval with high accuracy and interesting interaction interface by using python deep learning using tensorflow .

Here is list of what I can do with expertise for you -

- extract insights from social networks data from Facebook, Fitbit, Finance, Twitter, Pinterest, Youtube, Spotify, & more social networks data.
- create great charts or graphs using Tableau using python 3.7 library named tabpy.
- NLP model training for entity detection, NER, NLG or natural language generation, chatbot, question answering system and classification using transformer model such as by using python library transformer with hugging face transformer that has about 10 type of state-of-the-art pre-trained embeddings and models using python tensorflow for transfer learning on existing NLP dataset to fine tune it for better accuracy and performance. I can also use multi head attention, CNN, RNN, LSTM, Dense and hybrid models by combining these for NLP model building with high accuracy and better performance. I can also do training from scratch/using transfer learning/use state-of-the-art pre-trained vector embeddings such as BERT, xlnet, elmo and glove models using other python libraries.
- And do DIVERSE TEXT PROCESSING , Analysis and visualization, document classification, segmentation, topic modeling, intent extraction, NER extraction, NLU And NLG .
- BUILD ANY TYPE OF CHATBOT.
- POSE ESTIMATION, COMPUTER VISION BY TRANSFORMERS MODEL, image and video classification, segmentation, tracking, recognition, identification, watershed, yolo v3 training from scratch/using as it is,

training image processing models from scratch/using transfer learning/using state-of-the-art pre-trained computer vision models such as

mxnet,

inception v4,

efficientnet,

alexnet,

resnet- 32, 50, 56, 101,

imagenet,

open images,

&

inception resnet v3.

Also, I can convert any image classification model into yolo v3 like tracking project with state-of-the-art performance, tracking speed and tracker graphical design.

For predictive analytics data science projects, do with expertise -

- on excel/csv of about 1 GB, do

data cleaning,

missing data changing

&

wrangling.

- super hyperparameter tuning for deep learning models.

- use 16 types of evaluation metrics.

- save projects's data to csv, json, XML, text, pickle, hickle, hdf5 and save trained models to tflite for android applications and also save models to hdf5, ckpt & pb formats.

- data exploratory analysis and statistical trend extraction with lot of visualizations.

- automation of combination of individual or mix of machine learning and deep learning neural network models or algorithms to predict data with more accuracy.

- do feature extraction, engineering and importance on dataset to reduce the dataset and model size and find better accuracy of the trained models predictions.

I can automate download of data from online and extract text from voice data in audio or video using speech\_recognition python library with sphinx pre-trained model,

and can also use other trained models such google API use

or can train a new model with new dataset for better and/or domain specific speech recognition trained model.

I can train punctuation model using python code with new dataset

or use available pre-trained models with python library named deepcorrect.

I can also,

do speech synthesis using python library pyttsx3,

text translation,

text grammar and spelling correction and

deep sentiment analysis with vadersentiment/textblob/nltk/BERT

or

train a sentiment analysis model from scratch or train in addition to transfer learning from pre-trained models and text data pattern visualization using python 3.7 libraries like using wordcloud, nltk and many other text data visualization functions.

Also, if the dataset is too large to train in required time, for predictive analytics, NLP and computer vision projects ,

save the model checkpoints or last training stage

and later start with these last training state.

And, I can do transfer learning and automatic learning from ongoing interact data for local and cloud AI and/or data science applications using python 3.7 and cloud operate options.

Currently I work in python 3.7 in anaconda jupyter ipynb notebook by

tensorflow,

keras,

pytorch,

transformers for NLP, NLU and NLG (with Google's state-of-the-art transformer multi head attention model and the more recent Google brain research model - REFORMER built as modified transformer with more performance and prediction accuracy. And, with other transformers models, I implemented 10 models and can use many other hugging face transformer models, about 16+ state-of-the-art performance and speed pre-trained models, including BERT 3+ variations, elmo, xlnet and lot more pre-trained models for various NLP, NLU and NLG tasks). Also, out of tensorflow and pytorch for transformers, I use tensorflow as root layer of the transformers python library) ,

selenium,

scrapy,

regex,

speech\_recognition (speech recognition),

deepcorrect (text punctuation),

pyttsx3 (speech synthesis)

vadersentiment,

spacy,

nltk,

textblob,

rasa,

opencv,

yolo v3,

sklearn,

django,

flask,

tflearn,

automl by autokeras (super hyperparameter tuning of training model),

pyspark (big data),

MLib (big data),

tabpy (data visualization),

plotly (data visualization),

dash,

ggplot,

plotnine,

matplotlib,

seaborn,

wordcloud,

pyodbc (SQL and NoSql database connection using python as local or remote and do changes in databases),

SQLAlchemy,

and other python libraries.

I can automate database tasks using python pyodbc and sqlalchemy libraries to connect to SQL server, mongodb, postgresql, sqlite, mysql and jdbc and do copy, update with replace and get data from local or remote connection with better performance.

Link of my github AI code repository is-

[https://github.com/pritam-code-ai/Pritam\\_AI\\_code\\_in\\_python\\_1](https://github.com/pritam-code-ai/Pritam_AI_code_in_python_1)

.

My recently uploaded on 29 September, 2020 state-of-the-art AI projects are at github link -

[https://github.com/pritam-code-ai/commercial\\_type\\_NLP\\_computer\\_vision\\_analytics\\_projects](https://github.com/pritam-code-ai/commercial_type_NLP_computer_vision_analytics_projects)

.

And

Link of my kaggle data science projects is-

<https://www.kaggle.com/pritamcodeai/notebooks>

.

You can find lot of commercial and production level NLP and computer vision projects done by me in these links.

You can check these.

Also, I worked on huge datasets such as csv file of 4 million rows.

I think I can deliver you great data science projects using python 3.7 if you allow me in your team.

I am eagerly waiting for your reply.

Hope to hear from you soon.

Yours sincerely,

Pritam

## Work Experience

---

### **R&D Lead Data Scientist, computer vision engineer, NLP data scientist, data scientist**

Company - USA, Canada, Germany. freelancer- upwork and truelancer.

August 2018 to July 2020

Hi ,

I have data science projects clients from USA, Canada, Japan and india with their happiness on my completed work.

AS A MENTION,  
SOME OF MY IMPORTANT WORK EXPERIENCES ARE -

1.

I worked as co-founder and R&D CHIEF AI and top Manager on financial internet data parsing, NLP and prediction in an AI financial prediction solution startup company from USA of an angel investor , invested in multi companies. I leaded 10 interns from

""Harvard and MIT as masters and phd""

in this company and done multiple projects simultaneously for better performance and also done interviews to hire interns for this company. AND I LEARNED LOT OF MANAGEMENT SKILLS OF AI PRODUCT DEVELOPMENT FROM THE ANGEL INVESTOR AS A CO-FOUNDER.

2

I worked as lead AI Manager for interns for an big potential AI event marketplace startup from USA to build their recommendation system and lead their AI intensity team.

3

I worked as a R&D AI lead of NLP, NLU, NLG, computer vision and predictive analytics for a big potential online workpface and venue location marketplace startup from UK.

4

I worked for a geospatial solution company for satellite prediction of soil vegetation.

5

I worked for a German company on search technology and docker system improvement of the companies core product and done several fixes and provided interesting ideas with implementation option.

6

I worked for a supply chain AI automation company from USA and selected to lead 4 interns in this company.

7

I worked as an CHIEF AI and Lead AWS cloud engineer for financial fraud detection AI company from India with very big customers from USA and Africa.

8

I worked for 1 year in a big financial project as a portfolio optimizer program with dataset pre-processing to be ready for optimization on close price column and calculated log returns and shrinkage return for long short PORTFOLIO optimization , pre-filter by evolutionary algorithm and after that to select date range from start and end calender box, input typed days returns and risk free rate, select candidates selection metrics from about 33 hybrid and single metrics including treynor, ULCER INDEX, sterling, burke, alpha, beta, sharpe, systematic risk, unsystematic risk and many other and MENTIONABLE latest implementation of rachev version 1 and rachev percentile version and farinelli tibletti aggressive and conservative metric versions with hybrid on these from RECENT DEVELOPMENT RESEARCH PAPERS and then single and about 6 hybrid balancing as targets for efficient frontier optimization with visualization and result table with weights and about 8 metrics and these products as windows 10 GUI desktop application for by tkinter, pandastable, tkcalender, pyinstaller and lot of other python libraries as exe product for 1 of my client from USA .

9

I worked for 1 month in a covid 19 project and built a covid 19 x-ray image recognition ipynb file to py file to tkinter desktop GUI application using pyInstaller python library with lot of configuration to build exe compatible trained model file and the trained accuracy was 92 % and validation accuracy is 100 % with roc\_auc score more than 92 % and raw accuracy, specificity and sensitivity are each 100 % .

10

I was got invitation to build recommendation system by a company from Canada as co-founder position and by a company from India as CTO position.

## Education

---

### **Online Certificate in Computer vision python, NLP python, predictive analytics python, financial python, internet data scraping by python libraries and hyperledger blockchain .**

Udemy.com online

October 2018 to August 2019

### **Bachelor's in Computer Science and engineering**

CIT - west bengal, india

July 2013 to July 2017

### **CERTIFICATION in ADVANCED JAVA**

WEBTEK - Kolkata, West Bengal, India

## Skills / IT Skills

---

- Data Analysis (6 years)
- SQL (6 years)
- Excel (10+ years)
- Data science (6 years)
- Machine learning (6 years)
- Deep learning (6 years)
- Data visualization (6 years)
- Computer vision (6 years)
- Natural language processing (6 years)
- Tensorflow (6 years)
- Keras (4 years)
- NLTK (6 years)
- OpenCV (6 years)
- Classification (6 years)
- Clustering (6 years)
- Regression (6 years)
- AutoML by autokeras
- Recommendation engine (6 years)
- Google Dialogflow chatbot development (3 years)
- Scikit-learn (6 years)
- Ensemble (6 years)
- Hyper parameter tuning (6 years)
- Matplotlib (6 years)



- Seaborn (4 years)
- D3.js (3 years)
- YOLO v3 (4 years)
- internet data scraping (5 years)
- Big data by pyspark (4 years)
- Reinforcement learning (2 years)
- Dataset wrangling (6 years)
- Dataset exploratory analysis (6 years)
- Artificial intelligence (3 years)
- Neural network (6 years)
- Excel
- Business Intelligence (2 years)
- testing (2 years)
- HTML (3 years)
- Tableau (2 years)

## Online Profile

---

[https://github.com/pritam-code-ai/Pritam\\_AI\\_code\\_in\\_python\\_1](https://github.com/pritam-code-ai/Pritam_AI_code_in_python_1)

<https://www.kaggle.com/pritamcodeai/notebooks>