704/1, Bruchweisenanlage 4
66125 Saarbrucken
Gemany
+49 17687136243
deb_4u22@yahoo.com
debjitpaul.github.io/debjit.github.io

Debjit Paul

Experience

Oct.2016–Nov.2016 **Research Assistant**, *Saarland Univesity*, Saarland, Germany.
- **Department:** Foundation of Exact Algorithm
- Programming tasks for the organization of the programming challenge **PACE**.
- The goal is to generate instances to be used in the challenge.
- **Technology:** Python, Shell Scripting **Github Page**

Nov.2015–Feb.2016 **Research Assistant**, *Saarland Univesity*, Saarland, Germany.
- **Department:** Machine Learning Group
- Crafted Algorithms using Python to filter data (Arxiv dataset) collection results.
- Concept utilized are Natural Language Processing, Prepossessing text data.
- **Technology:** Python, Shell Scripting.

April 2015–July 2015 **Student Research Assistant**, *Saarland University, DFKI*, Saarland, Germany.
- **Department:** Information and Services Systems
- Developing features in OntoUML.
- Designing Information Systems based on Conceptual Modeling.
- **Technology:** Java, Protege, OntoUML.

Education

2014–2017 **MSc. Computer Science**, *Saarland University*, Saarbrücken,Germany, *Grade-Yet to be given.*

Thesis Multitasking Learning with unreliable labels

Supervisor Prof.Dr.Dietrich Klakow, Head of Language Spoken System Group

Description In recent times neural networks are producing impressive results due to the presence of large data sets. However, a common known problem in classification task is unreliable labels, due to artificial annotators, human annotation mistakes. With the increase in size of training data, neural network can outperform several classical Machine Learning algorithms. This thesis explored how to handle un-reliable labels for Big data and how to clean the data while training a Single task Neural network and also Multi-task Neural network.

Concept RNN(BiLSTMs), Expectation Maximization Algorithm

2010–2014 **B.Tech in Computer Science**, *GuruNanak Institute of Technology*, Kolkata,India, *Grade-8.73/10.*

Thesis Improved Algorithm for Human and non Human Object Detection

Selected Project

Topic Machine-Learning-Tools-on-Iris-Dataset

Description This tool is made to perform Machine Learning classification on Iris Dataset. Machine Learning classification tools are as follows: Support Vector Machine, Logistic Regression, Adaboost, Random forest. **Github Page**

Topic Pre-Processing and NLP Tagger Tool

Description Transforming a unstructured text data into CONLL format NLP tagging file. This tool can handle large unstructured data. Then can be used to tag words using artificial taggers such as Senna or Brill Tagger. **Github Page**

Topic	Bidirectional long short-term memory
Description	In recent times Bidirectional long short-term memory (BiLSTMs) networks has proven success for several NLP sequence task such as POS-tagging, NER tagging and Chunking. Implemented a BiLSTMs tagger with word and unicode byte embeddings.

Technical skills

Language	Python, R, Shell Scripting,Java, C,Matlab, MySQL
Concepts	Machine Learning , Statistical Learning, Data Mining
Tools	Pycharm, word2vec, Tensorflow, sckit-learn, Dynet, Standford CoreNLP, Theano, Eclipse, Matlab, Numpy

Course Work

Machine Learning, Neural Networking, Artificial Intelligence, Statistical Learning, Convex Analysis and Optimization, Data Mining and Matrices, Topics Algorithm in Data analysis, Statistical Natural Language Processing, Image Processing and Computer Vision

Publication

Title	Anup Kumar Thander, Goutam Mandal, Debjit Paul **"Numerical Comparison of multi-step iterative methods for finding roots of non-linear equations, International Journal of Mathematics Trends and Technology "**, Volume4 Issue 8-September 2013 [ISSN: 2231-5373]

Languages

English	**Proficient**
German	**Basic**
Bengali & Hindi	**Native**

Hobbies

- Playing Football, Cricket and Tennis
- Photography

References

- Available upon request

Declaration

- I hereby declare that all the details furnished above are true to the best of my knowledge and belief