Manasa Mohan Pawar

Nationality: **Date of birth:** 02/06/1998

📞 **Phone number:**

✉ **Email address:** _____

in **LinkedIn:** _____

📍 **Home:**

ABOUT ME

Goal-Oriented, Facilitator, Skilled in programming languages, and a budding data scientist.

Interested in data analytics, motivated to pursue a career as a Data Scientist to become an expert in applying Machine Learning algorithms to various real-world problems for automating processes and bringing awareness to data-driven decision-making.

Primary skills and interests:
• Data Analysis and Performance Reporting
• Teamwork and Cross-functional Communication
• Problem Solving and Process Improvement
• Multitasking and Time Management

WORK EXPERIENCE

Research Assistant
Consiglio Nazionale delle Ricerche,CNR [22/05/2022 – Current]

Company area: Italian public research body
Institute of Applied Mathematics and Information Technologies "Enrico Magenes"-IMATI Sezione di Genova -Via de Marini, 6, 16149 Genova www.imati.cnr.it

Supervisor/ PI : Antonella Galizia (CNR, IMATI, Genova)
Under her supervision and guidance, I am currently working and exploring the world of environmental science, meteorology weather, and climate to understand the problem and propose solutions with the help of artificial intelligence and data science.

Activity:
I conducted research and studied state-of-the-art related to the data processing and knowledge extraction of climate change and its impact in association with the I-CHANGE project which intends to promote a learning process aimed at improving citizens' knowledge of the phenomenon of climate change and to make them understand how to mitigate its impacts through behavioral changes and/or structural actions (such as nature-based solutions). Few of the responsibilities and activity are as follows:
  • Gathered and analyzed data from a variety of sources, including satellite imagery, earth observations, geographic data, and sensor readings to estimate the impact of climate change.
  • I had the opportunity to attend some workshops hosted by renowned European projects where I learned about the latest trends of data science in the field of meteorology and environmental science.
  • Enriched my knowledge by learning the tools and techniques used to support my understanding and explore climate data.
Along with the I-CHANGE project, I also had the opportunity to research biomedical data(HiC data) to cluster cells and find patterns between healthy and unhealthy cells under the supervision of Claudia Caudai - Pisa - CNR-ISTI

Data scientist
ReLOG3P S.R.L. [02/2022 – 05/2022]

Company area: Logistics and Information Technology
Salita Nuova Di N.S Del Monte 6A/29, 16143, Genova https://relog3p.com

Soon after my graduation, I started working at ReLOG3P which is an Innovative Startup that provides high-tech services and products with the aim of supporting the world of logistics (end-to-end) to make its contribution to achieving the Sustainable Development Goals set by the United Nations (Agendas, mainly through a responsible and ethical use ("Responsible and Ethic Research & Innovation") of the typical technologies of the 4th Industrial Revolution (AI, Data Science, Blockchain, quantum computing, etc.)

Activity:
Since it was a start-up I had to "wear different hats" and carry out some of the activities as follows:
- Gained experience in querying databases and using statistical computer languages (R, Python, SQL, etc.) to manipulate data and draw insights from large data sets.
- Gained experience working with and creating data architectures.
- Improved knowledge of a variety of Machine Learning techniques (Clustering, Decision Tree Learning, Artificial Neural Networks, etc.) and their real-world applications.
- Increased familiarity and knowledge with, proposed, and applied the existing and upcoming Ethics and Responsible Innovation, Privacy, and Security by Design regulations, standards, and best practices and solutions.
- Assessed the convenience, viability, effectiveness, and accuracy of new data sources and data-gathering techniques.
- Data mining and analysis of data from internal and external databases to drive optimization and improvement of product development, marketing, and business strategy.

Machine learning Intern
Livewire [01/2019 – 02/2019]

Company area: Education, Training, Research, and Development
Channasandra Main Road, Whitefield, Bengaluru, 560066 www.livewireindia.com

Towards the end of my bachelor's degree, I became a machine learning intern at Livewire. This is when my passion for data science sparked. After this, I decided to perceive my master's in the same field.

Activity:
While I worked, I had my first encounter with Machine Learning algorithms and Big data. Specifically,
- I gained experience working on real-world datasets in order to apply skills like machine learning, big data, and neural networks.
- Knowledge in the field of Computer vision, Distance metrics, Time series analysis, Scikit-learn, Python, and Tensor Flow.

IT Intern
Hindustan Aeronautics Limited [08/2018 – 09/2018]

Company area:- Aeronautics, Aerospace, Naval.
443, 1st, A Cross Rd, LBS Nagar, Kaggadasapura, Bengaluru, Karnataka, 560017 https://hal-india.co.in/

Hindustan Aeronautics Limited is an Indian state-owned aerospace and defense company, headquartered in Bangalore, India. HAL is one of the oldest and largest aerospace and defense manufacturers in the world today. During my bachelor's study, I got selected for an internship that focused on the Information Technology department of the RWRDC branch of HAL.

Activity:
- I worked alongside the employees of the IT dept. and exercised my programming skills in .NET and developed a job portal for new applicants applying for HAL.
- During the internship, I had the opportunity to experience the server rooms of high-performance computers.

- I had the chance to better understand the theoretical concepts of application development, and computer networks and gained my first experience working with a real-time application.

EDUCATION AND TRAINING

Master's Degree in Data Science
Amrita Vishwa Vidyapeetham [2019 – 2021]

Final grade: 8.2/10

Amrita Nagar, Choodasandra, Junnasandra, Bengaluru, Karnataka,India 560035
https://www.amrita.edu/

Thesis Title: Multi-layer, multi-modal medical image intelligent fusion https://link.springer.com/article/10.1007/s11042-022-13482-y

Thesis description: This thesis is also published in a journal affiliated to Amrita Vishwa Vidyapeetham. In this work, we proposes a novel multi-layer, multi-tier system called Multi-Layer Intelligent Image Fusion(MLIIF) with deep learning(DL) networks for visually enhanced medical images through fusion. Implemented deep feature-based multilayer fusion strategy for both high-frequency and low-frequency components to obtain a more informative fused image from the source image sets.

Supervisor: Dr. Tripty Singh - Amrita Vishwa Vidyapeetham

Courses:
-Linear Algebra and Optimization Techniques
-Data Structures and Algorithms
-Big Data Mining
-Statistical Learning
-Computational Intelligence
-Data Preparation and Analysis
-Machine Learning
-Time Series Analysis and Forecasting
-Deep learning

Master's in Computer Science
Università degli Studi dell'Aquila [2020 – 2021]

Final grade: 110/110

Palazzo Camponeschi, Piazza Santa Margherita, 2,67100, L'Aquila,Italy
https://www.univaq.it/

Thesis Title: Modified data aggregation approach to improve the energy efficiency using Artificial Intelligence

Thesis description: Wireless sensor networks communicate via sensor nodes that are battery-powered. Hence overgrowing data with limited resources need to be managed. Work focuses on the importance of data aggregation to achieve a data transmission rate by applying artificial intelligence using python.

Supervisor: Matteo Spezialetti,Filippo Mignosi - DISIM, Univaq

Courses:
- Software Architectures
- Service Oriented Software Engineering
- Information and Network Security
- Software Engineering for Autonomous Systems
- Information Retrieval
- Research topics in Software Architecture

Bachelor of engineering in Computer Science & Engineering
Gopalan College of Engineering and Management [2014 – 2019]

Final grade: 8.15/10

Basavanagar, 181/1, 182/1, Seetharampalya - Hoodi Rd, Behind SAP Labs, Hoodi, Bengaluru,560048, Karnataka,India https://www.gopalancolleges.com/gcem/

Thesis Title: Semi-Supervised Video-to-Image Adaptation for Human Action Recognition

Thesis description: Human Action Recognition using Machine Learning algorithms and a Digital Image Processing Technique like OpenCV was used to split video frames on video for the action of people detection in given frames.

Supervisor: Dr.A.A. Powly Thomas (Principal- GCEM)

SPECIALIZATIONS / CERTIFICATIONS

Deep Learning Hands-On
[12/07/2022 – 20/07/2022]

Università Genova,2022

It covered the basic principles of computer vision and visual perception in artificial agents, including theoretical classes, application examples, hand-on activities.

20 hours course.

Link: https://malga.unige.it/education/schools/cvcc2022/

Introduction to Data Analytics for Business
[10/2021 – 10/2021]

Coursera

This course covered the key areas as the analytical process, how data is created, stored, accessed, and how the organization works with data and creates the environment in which analytics can flourish.

12 hours course.

Link: https://www.coursera.org/account/accomplishments/verify/GNBJLKUL438M

AI for everyone
[05/2020 – 05/2020]

Coursera

- The meaning behind common AI terminology, including neural networks, machine learning, deep learning, and data science

- What AI realistically can--and cannot--do

- How to spot opportunities to apply AI to problems in your own organization

- What it feels like to build machine learning and data science projects

- How to work with an AI team and build an AI strategy in your company

- How to navigate ethical and societal discussions surrounding AI

11 hours coursework

Link: https://www.coursera.org/account/accomplishments/certificate/LEB22LP2XC6Y

Java/SQL development
[01/2019 – 03/2019]

Qspiders

It is a training institute, which I was entitled to get free training after qualifying the interview. Apart from Java development and SQL, mock interview and additional placement activity was also organized.

It was 2 months long.

Link: https://qspiders.com/batches

Python Specialization

[02/2018 – 09/2018]

Coursera

Introduced fundamental programming concepts including data structures, networked application program interfaces, and databases, using the Python programming language. To use the technologies learned throughout the Specialization to design and create own applications for data retrieval, processing, and visualization.

8 months specialization course

Link: https://www.coursera.org/account/accomplishments/specialization/LUQ7XPJMPHXM

Ethical Hacking Workshop

[2017]

Indian Institute of Science

This workshop covered a variety of topics such as Scanning a Network and Finding Out Vulnerabilities, Hacking a Windows System, Hacking a Linux system, Social Media Hacking, Google Hacking Database, DDoS and many other.

It was a 2 day workshop

Link: https://www.pravega.org/files/workshops/EthicalHacking-Offensive.pdf

LANGUAGE SKILLS

Mother tongue(s): **Marathi**

Other language(s):

English

LISTENING C2 **READING** C2 **WRITING** C2
SPOKEN PRODUCTION C2 **SPOKEN INTERACTION** C2

Italian

LISTENING B1 **READING** A2 **WRITING** A2
SPOKEN PRODUCTION A2 **SPOKEN INTERACTION** A2

Hindi

LISTENING C2 **READING** C2 **WRITING** C2
SPOKEN PRODUCTION C2 **SPOKEN INTERACTION** C2

DIGITAL SKILLS

Tools and Technology

MySQL / Web Scrapping / Statistical Data analysis / django / T ensorflow / Keras Framework / Oracle SQL / HTML5/ CSS, Javascript / Modelling software (Weka) / Computer Vision / GitHub / Machine Learning / Basics of React.js / Power BI/ Power Query / Docker / Artificial Neural Networking

Programming languages

Python / R / Matlab / Java / C

Industrial Knowledge

Time Series analysis / Business analysis / Cloud computing (Google Cloud, IBM Cloud) / Data Gathering, Data Processing, Data Visualization, Data Analysis / Microservice architecture REST API / object-oriented programming / Data structures / Service oriented (SOA)

DIGITAL SKILL DESCRIPTION

Web scraping

As we are mainly depending on sites like Kaggle to get a dataset for data science training, in real-time it is not possible to get the data set ready, for that it is mandatory to learn web scrapping. For Web scrapping, learned python tools Beautiful Soup, Scrapy, and urllib.

Data Visualisation

For understanding data, I learned tools like Tableau, sweetviz and dtale. All the projects I have done to date required certain phases of Data Cleaning, Data Aggregation, model fitting, model training, and helped me in acquiring exploratory analysis skills. For exploratory analysis, learned certain visualization tools in python e.g., Matplotlib, Seaborn, and plotly.

Deep Learning

By applying Deep Learning algorithms e.g., CNN, LSTM, RNN, I worked on projects like Eye Gaze, Captcha Reader, Image Caption Generator, and Credit Card Analysis by also apply image processing. For getting a better understanding of certain machine learning algorithms e.g. OrthogonalMatchingPursiut, RANSACRegressor, ElasticNetCV, Lasso, I worked on a project to check the performance on the same dataset.

In the medical field, I worked on several types of datasets like Parkinson's, Diabetic patient's data set, and Heart patient's data. To predict several factors. In the wrap during working on all these projects, I worked with the following python tools: Pandas, Numpy, Matplotlib, Seaborn, Keras, Tensorflow, Wordcloud, Textblob, Pytorch, Scikit, Scipy, StatsModel, Plotly.

Machine learning

In the aspect of Machine learning I have worked on the following algorithms :

Linear Regression, Logistic Regression, Decision Tree, SVM, Naive Bayes, kNN, K-Means, Random Forest, Dimensionality Reduction Algorithm, Gradient Boosting Algorithm: GBM XGBoost LightGBM CatBoost

For the deep learning projects, I have worked on the following algorithms :

Convolutional Neural Networks, Long Short-Term Memory Networks, Recurrent Neural Networks, Generative Adversarial Networks, Multilayer Perceptrons, Self-Organizing Maps, Deep Belief Networks

Big data modelling

In the project flight data analysis, I worked on Big data modelling algorithm MapReduce to make the data usable and used it to analyze the frequency of particular flights. In the aspect of Big data modelling, I have gathered conceptual knowledge of the following algorithms :

Map Reduce, Regression Trees, and the tools used for big data modelling were Hadoop tools Hive and pig.

Database Management

While working on a project based on web development I worked with MySQL and Handled all database-related duties, which required deep knowledge of SQL language and also a thorough knowledge of JDBC and ODBC was acquired. Also worked on Oracle 11g.

PROJECTS

Movie Database Application

[08/2021 – 08/2021]

During an application for a research project in University of L'Aquila, Under the guidance of De Gasperis Giovanni -(Univaq) I deployed a movie database with flexible query search and attractive user interface.

Tools: Django framework, Docker environment.

L'Aquila Railways

[10/2020 – 02/2021]

During my coursework in University of L'Aquila, for the course Software Engineering for Autonomous Systems I developed this project. The goal of this project is to develop a web application with an elegant and easy to use design. For the user to check the availability of trains and seats. And also book the ticket.

Tools: Eclipse, Maven, CXF, MySQL Workbench, React, REST API, SOAP API

Eye Gaze

[10/2020 – 11/2020]

This was part of deep learning course at Amrita. A model to train and estimate the position of the user's eye captured from camera as an input, to get the angle of the user's line of sight as an output.

Tools: Google Colab, Keras, Scipy, Scikit, Matplotlib, Pandas, NumPy, Seaborn, itertools, Glob, TQDM

Algorithms: Convolutional Neural Network

Dataset: Dataset is imported from Kaggle, it contains a total of 213,659 images from 15 participants. The number of images collected by each participant varied from 34,745 to 1,498.

Credit Card Analysis

[07/2020 – 08/2020]

This project was developed for the course machine learning in amrita to demonstrate Regularization, this model was trained with various Optimizers and loss functions to improve the accuracy of data provided of credit transfers from the credit card.

Tools: Google Colab, Keras, Scipy, Scikit, Matplotlib, Pandas, NumPy, Seaborn,

Regularizer.

Algorithms: Convolutional Neural Network, LinearRegression, Ridge, Lasso

Dataset: The dataset is imported from Kaggle, it contains transactions made by credit cards in September 2013 by European cardholders. This dataset presents transactions that occurred in two days, where we have 492 frauds out of 284,807 transactions. The dataset is highly unbalanced, the positive class (frauds) account for 0.172% of all transactions.

Accident Analysis

[05/2020 – 06/2020]

For time series analysis coursework I developed this project, a model trained to predict the number of accidents that might happen in a region. This model was trained using seasonal ARIMA, Time series analysis and checking accuracy by MSE.

Tools: Google Colab, Keras, Scipy, Scikit, Matplotlib, Pandas, NumPy, Seaborn, itertools, statsmodels

Algorithms: ARIMA

Dataset: This dataset is imported from Kaggle, it provides detailed road safety data about the circumstances of personal injury road accidents in GB from 2014 to 2017. The statistics relate only to personal injury accidents on public roads that are reported to the police and subsequently recorded, using the STATS19 accident reporting form.

Parkinson's Disease Data Preparation and Analysis

[03/2020 – 04/2020]

For the course work, Data preparation and analysis this project was developed. Various statistical models were demonstrated to infer the best statistical model among them, which predicts with high accuracy using python libraries. A similar work was done using R.

Tools: Google Colab, Keras, Scipy, Scikit, Matplotlib, Pandas, NumPy, Seaborn

Algorithms: Logistic Regression, KFold

Dataset: The data used in this study is imported from Kaggle, details of this dataset were gathered from 188 patients with PD (107 men and 81 women) with ages ranging from 33 to 87.

Stock Price Prediction

[02/2020 – 03/2020]

This is one of my personal projects I worked to understand the concepts during my master's. This project is about using LSTM recurrent neural networks in open, high, low and closing prices of Apple Inc. stocks value Prediction. It includes two sequential LSTM layers that have been stacked together and one dense layer that is used to build the RNN model using Keras deep learning library.

Tools: Google Colab, Keras, Scipy, Scikit, Matplotlib, Pandas, NumPy, Seaborn, Tensorflow

Algorithms: Convolutional Neural Network, Long Short Time Memory, Recurrent Neural Network

Dataset: The dataset is taken from yahoo finance's website in CSV format. The dataset consists of Open, High, Low and Closing Prices of Apple Inc. stocks from 3rd January 2011 to 13th August 2017 - a total of 1664 rows.

Flight Delay Analysis

[02/2020 – 03/2020]

In my course work in Amrita, under the course big data I developed this project using MapReduce on a large dataset, for counting the frequency of canceled flights for a particular year using map-reduce framework and JAVA.

05-12-2022