Santosh Kumar

New Delhi, India
+91-9988882597 santosh.iitk7@gmail.com www.linkedin.com/in/santosh-kumar-iitk18

Summary

Dynamic Data Scientist and Computational Engineer versed in applying machine learning models and advanced data analytics to enhance business solutions and operational efficiencies. Expertise in Python, SQL, and collaboration with cross-functional teams has led to noteworthy accomplishments such as boosting demand forecasting accuracy and optimizing delivery routes. Eager to harness these skills to drive impactful data strategies in future projects.

EXPERIENCE & POSITION OF RESPONSIBILITY

Hon Hai Foxconn India pvt. ltd **Jan 2024 - June 2025**
Computational Engineer *Bangalore*

- Analyzed complex challenges and implemented effective solutions, enhancing project outcomes in dynamic environments
- Applied AI/ML techniques using Python and TensorFlow for predictive model testing and optimization, improving regression task accuracy
- Modeled and analyzed data to create analytical models, optimizing system performance and reducing errors
- Utilized algorithms and data structures to efficiently solve engineering challenges, improving computational processes
- Handled large datasets using Hadoop and Spark for advanced analysis, enhancing data processing capabilities
- Worked Cross-Functional Collaboration with international teams to clarify customer requirements and deliver innovative solutions.
- Collaborated with overseas clients using Python and Shell scripting to bridge technical gaps between vendors and customers, improving communication and project outcomes
- Developed scripts using Python and PySpark to automate testing processes, improving workflow efficiency and reducing manual errors
- Applied proficiency in calculus, linear algebra, and numerical methods to support computational modeling, enhancing the accuracy and reliability of engineering simulations
- Implemented energy-efficient computing practices using TensorFlow and Keras to align with green engineering initiatives, reducing energy consumption in computational processes
- Utilized Linux/Unix systems and common commands to streamline system operations, enhancing efficiency and reliability

Oms logistics **Sep 2023 - Dec 2023**
Data Scientist *Pune, India*

- Leveraged machine learning, statistical modeling, and data visualization to improve operational efficiency, reduce costs, and enhance customer satisfaction in logistics and supply chain optimization.
- Developed predictive models to forecast demand and optimize inventory levels, reducing stockouts by 25% and overstocking by 10%.
- Utilized geospatial analysis and machine learning algorithms to optimize delivery routes, resulting in a 10% reduction in fuel consumption and a 15% decrease in delivery times.
- Managed large datasets using MySQL databases to track shipments and inventory across multiple warehouses and implemented machine learning techniques and statistical methods to predict and analyze vendor performance and potential supply chain disruptions, enabling proactive mitigation strategies
- Worked closely with logistics teams to implement data-driven solutions, improving operational efficiency by 25%.

Adanac Digital **Jan 2021 - Aug 2023**
Data Analyst *Mumbai, India*

- Developed and deployed predictive sales models using Python (Pandas, Scikit-learn), improving forecast accuracy by 20% over previous benchmarks.
- Analyzed historical sales data across multiple regions and product categories to identify trends and seasonality, enabling data-driven demand planning.
- Designed and automated interactive dashboards in Tableau/Advance Excel to visualize sales forecasts and performance variance across business units.
- Collaborated with marketing and finance teams to align sales predictions with promotional campaigns, reducing stockouts by 10% and excess inventory by 8%.
- Conducted A/B testing and time-series modeling (ARIMA, Prophet) to evaluate pricing and promotional impact on future sales trends.
- Delivered monthly executive reports summarizing forecast accuracy, business risks, and actionable insights, contributing to increase in forecast reliability.

Tutorpoint **June 2020 - Jan 2021**
Thermodynamics Subject Matter Expert
Six months working as Thermodynamics subject matter expert in Tutorpoint.

- Conducted one-on-one sessions to explain the laws of thermodynamics, resulting in a 40% improvement in student exam scores by clarifying challenging concepts like entropy and enthalpy
- Designed tailored lesson plans to address individual learning needs, focusing on thermodynamic processes such as isothermal and adiabatic changes, leading to measurable academic progress.
- Provided academic coaching that resulted in a 90% pass rate among students struggling with advanced topics like transient energy systems

Indian Institute of Technology, Kanpur **July 2018 - May 2020**
Two years assigned the responsibility of Teaching Assistant in Aerospace Department, IIT Kanpur.

TECHNICAL SKILLS AND RELEVANT COURSES

- **Programming Skills**: Python, Advance Excel, R-Language, Fortran, NLP, SQL, MS-Office 365, Shell, Perl, Linux Operating System, Excel-VBA
- **Development Skills**: Data Science & AI, Generative AI, Machine Learning Algorithm, LLMs, Natural Language Processing, Data Modelling, ETL, EDA Analysis, Deep Learning, Web Scrabbing, Regression, Statistical Modelling, Object Detection, ChatGPT, Clustering, Cloud Computing, Fraud Detection, Risk Management, Product Management, Object Oriented Programming, GPU programming
- **Visualization Tool**: PowerBI, Tableau, Advance Excel, Latex, JMP Software
- **Library**: NumPy, Pandas, Matplotlib, TensorFlow, SciKit-Learn, Pytorch, SciPy, Pyspark, spacy, Seaborn, OpenCV, Kreas, Cloud Computing, XGBoost

Projects

NLP Web Scraping

- Performed automated web scraping from XML files to extract large-scale textual data using Python and BeautifulSoup.
- Implemented data preprocessing pipelines including HTML parsing, stopword removal, lemmatization, and noise reduction using NLTK and SpaCy.
- Combined and cleaned text data from multiple XML articles to build a consolidated corpus for further analysis.
- Applied advanced text normalization and tokenization techniques for preparing data suitable for machine learning models.
- Developed and visualized word clouds and frequency distributions to identify key terms and patterns in text datasets.
- Engineered feature extraction techniques using CountVectorizer and TF-IDF transformers with normalization.
- Performed K-Means clustering to group similar articles and identify thematic clusters in both unigram and bigram (n-gram) representations.
- Conducted clustering evaluation using distortion plots (Elbow Method) to determine optimal cluster counts.
- Calculated text scoring mechanisms to rank articles based on word and sentence importance metrics.
- Built dynamic visual word cloud generation for each cluster group, aiding interpretation of clustering results.
- Demonstrated proficiency in handling end-to-end Natural Language Processing pipelines and unsupervised text clustering.

Color Detection Project (Computer Vision using OpenCV)

- Developed a real-time color detection system using OpenCV to identify and isolate red, green, and blue objects from live webcam feed.
- Designed color masks in HSV color space for accurate color segmentation and object detection under varying lighting conditions.
- Implemented efficient frame processing and bitwise operations to filter and display targeted colors in real time.
- Built robust scripts to detect every color except white, demonstrating flexibility in dynamic masking techniques.
- Gained practical experience in computer vision preprocessing techniques and color threshold tuning for optimized accuracy.

CERTIFICATIONS

- **Fundamentals to Become a Machine Learning Engineer on LinkedIn**:
  1. Artificial Intelligence Foundations: Thinking Machines
  2.Artificial Intelligence Foundations: Machine Learning
  3.Machine Learning with Python: Foundations
  4.Deep Learning: Model Optimization and Tuning
  5.Reinforcement Learning Foundations
  6. Generative AI: Introduction to Large Language Models
- **Data Analyst Certification**: Certified in Data Analytics Virtual internship from Trainity platform.
  Certified in 8 weeks data analytics specialization training from Trainity.
- **Product Management Certification**: Certified in Product management from Great Learning.

EDUCTIONAL QUALIFICATIONS

Indian Institute of Technology, Kanpur **2018 - 2020**
M-Tech, Aerodynamics

- **GPA:** 7.25 /10

Rashtrasant Tukadoji Maharaj Nagpur University (RTMNU) **2012 - 2016**
BE, Mechanical

- **GPA:** 8.6/10 (Distinction)

SCHOLASTIC ACHIEVEMENTS

- **GATE Score**: Achieved a GATE score securing an All-India Rank of 1250 in 2018 in Mechanical Paper, demonstrating strong analytical and problem-solving skills
- **3rd University Rank**: Secured 3rd university rank in third semester of Mechanical Engineering in RTMNU. (2014)
- **2nd Position in HORIZON-2013**: Secured 2nd position among of 2000 students in national level "HORIZON-2013", conducted by PCE NAGPUR. (2013)
- **Branch Topper**: 6 times branch topper in mechanical Engineering in college.
- **3rd Prize in Project Exhibition**: Got 3rd prize in project exhibition in intra-college in Nagpur. (2016)