CURRICULUM VITAE

Saurabh Chandra Keshari sckeshari344@gmail.com
+91-7540031218

Profile Summary:

- 2.5 years of experience in **Data Science.**
- Certified Expertise in **Machine Learning and Artificial Intelligence (AIML)** Program from **IIITH, Hyderabad**.
- Good knowledge in **Classical Machine Learning and Deep Learning**.
- Good understanding and practical knowledge of Pyspark and Apache Spark.
- Working with huge amount of time-series based sensors data both from SCADA and CMS systems.
- Certified in Python fundamentals by Pluralsight.
- Worked with popular libraries like SKLearn, NumPy, Matplotlib, Pandas, Plotly, Tensorflow, Seaborn, OpenCV, NLTK, etc.
- Worked on Flask framework in Python.
- Good knowledge of Data wrangling, Exploratory Data Analysis (EDA), Dimensionality reduction techniques and Feature Engineering.
- Worked on Classification, Regression, Clustering and NLP problems including several Kaggle competition. Link of my Kaggle profile below: https://www.kaggle.com/sckeshari344
- Hands-on End-To-End experience on ML pipeline.
- Actively involved in participation of daily/weekly project status meetings with Team Members, Team Lead and Higher Management.
- Quick learner and excellent team player having ability to meet tight deadlines and work under pressure
- Passion to read about new technologies and concepts.

Technical Skills:

Programming Languages	:	Python, C#
Operating Systems	:	Windows all versions
Python IDE for ML	:	Spyder, Jupyter Notebook, Kaggle kernel, Google Colab
Bigdata Platform	:	Databricks

EDUCATION:

- B.Tech (Bachelor of Technology) from SRMSCET, Bareilly, UP

Experience Summary:

- Currently working as Data Science Engineer in Utopus Insights (A renewable energy based company) fromMay'19 till date..
- Worked as a Machine Learning Engineer in **Evry India Pvt Ltd** (CMMi Level 5 Company)**,** Bangalore from December 2017 to till april'19.
- Worked as a Software Engineer in **Robert Bosch,** Bangalore from September 2014 to December 2017

Certifications:

- Certified Expertise in **Machine Learning and Artificial Intelligence (AIML)** Program from **IIITH, Hyderabad**.
- Machine Learning A-Z: Hands-On Python & R in Data Science, Udemy
- Docker for Data Science.
- Neural network and Deep Learning, Coursera
- Python Foundation certified, Plural Sight
- ISTQB (International Software Testing Qualifications Board) Certified
- Certified in ITIL Foundation Level.

Technical Work experience

Project #1: Scipher Rx (A predictive maintenance product)

Tool used: Databricks platform for big data analytics
Language used: Python and pyspark, Spark SQL, SQL
Period: May'19 – On going

- Scipher Rx is a predictive maintenance product of Utopus Insights used for assisting and predicting failure of Wind Turbines **ahead of time** based on time-series based SCADA and CMS (Vibration) data using various approaches and modelling techniques.
- Regional and Global based models are designed based on model type, controller type and rotor diameter.
- Farm based wind turbine analytical approaches are applied such as CDC, Efficacy, Normalization, unsupervised, etc to study each wind farms to address pain points of wind farm owners and also to advice on faulty/unhealthy wind turbines.
- Visualizing data to understand the signature/behavior of signals before various failure modes (service orders) such as Generator replacement, Generator bearing replacements, Gearbox replacement, pitch misalignment, sub-failure modes, etc.
- Understanding of domain plays a key role in order to optimize and understand every nuts and bolts of wind turbines for analysis and modelling.

Role and Responsibilities–Data Science Engineer

- Building models for various wind turbine failure modes such as Generator, Gearbox, main bearing, etc which case be global, regional and farm-based model.
- Proper selection of healthy and failure data for training-testing.
- Applying various technical approaches to understand complex sensors data.
- In-depth analysis of sensor data to get bring better insights from data using visualization and other statistical techniques.
- Understanding turbine behavior over time and alert farm managers about unhealthy turbines and their root cause.
- Proper cluster management and job schedule with minimal configuration setting of worker nodes for cost optimization.
- Attending daily scrum calls with lead and higher management.
- Understanding Gearbox and Generator bearings life cycle and various factors affecting components leading to detoriation such as improper Lubrication, misalignment, Inner-outer race damages, etc and technically communicating it to our customers using plots and statistical techniques.
- Enhancement in our present model and Improving feature selection techniques

Project #2: Next Best Retail

Period: Dec 2018 -Feb 2019

- How can we devise localized advertising strategies to increase retail store walk-ins with the same ease as leveraging online demographics for advertising content delivery in Facebook via Google Ads.
- Online analytics using DJI Tello Drone with IOT device mounted on it to transfer the data to cloud.
- Object detection using TensorFlow and deploying it on Drone.
- Offline analytics using Azure cognitive services such as Video Indexer.

Detect –

- Demographics (Age group, gender),
- clothing style (cultural recognition, appearance), (Jeans / traditional / party attire etc.)
- location / proximity (shop -character recognition), sentiment, group type (family / friends/ kids)
- activity recognition (partying, having dinner, playing snooker / ball, savoring beer, etc.),
- Clustering Shoppers - like Frugal shoppers / Enthusiastic shoppers etc.

Project #3: Dent Identification

Period: Nov 2018 – On going

- Extracting video of Car into unique frames.
- Building Deep learning model using CNNs (Mask R-CNN) to do segmentations for dent/damages in images.
- Comparing dent/damage images with CNN Siamese network.

Project #4

Title	:	Resume Ranking
Skill Used	:	NLP, String Matching and Ranking algorithm
Role & Team Size	:	Data Scientist & 4 Members
Tool used	:	Spyder
Period	:	Aug 2018- Oct 2018

Responsibilities:

- Understanding business scenarios to sort resumes based on HR processes based on certain parameters.
- Worked on creating end to end solution using Flask Framework in Python.
- Formulating basic architecture for resume's selection based on parameters such as skills, Year of experience, Qualification, etc. Assigning needed weightage to each based on requirement which is configurable.
- Reading resume of all formats using textract and parsing it. Used Tokenization, lemmatization, stemming, summarization and string-matching concepts of NLP.
- Inputs are raw Job Description and resume.
- Extracted phone no, name, email ids, any other links, experience, technical skills, non-tech skills, and final score based on scores from Experience, Skills and Qualification.
- Working on extracting skills & projects and integrating more features to the model such as Linked in and Github profile.

Project #5: Reduce Loan Mortgage
Period: June 2018 – July 2018

- Built an end to end Reduce Loan Mortgage application using Flask Framework, Python.
- Predicting whether to approve loan based on historical data of users.
- Formulating business rule based on Decision rules generated.
- Used XGBoost classifier for making prediction.
- Selection of features done based on EDA, manual selection and feature engineering.

Project #6: Chat bot

Period: Feb 2018 – April 2018

- Chat bot which accepts user input and replies with the understanding through entity recognition on the input text.
- Used FAQ as a knowledge base and formulating questions around it.
- Multiple questions formed for single intent and a single reply to it.
- Entity recognition and text cleaning done using NLP.

Project #6

Project Name	**: Gecko Information Systems INC**
Client	**:** Norway
Environment	**: .**Net, SQL, Oracle, Outlook, share point
Testing Tools	**:** Coded UI VSTS Automation, TFS, MTM
Period	**:** Feb 2016 – Jan 2018
Team size	**:** 8

Description:
The main product is the document, case and filing solutions for the public as negotiated by BOSCH. Gecko has over time become one of the leading development environments for this type of product. Over 350 Norwegian municipalities, private companies and government agencies - currently use our system. Apart from Norway, Sweden and Denmark municipalities are also using this product.

Products:

1. **ePhorte Elements**:
2. **ePhorte Web**:
3. **ePhorte SharePoint**:
4. **ePhorte Outlook**:

Responsibilities:

- ✓ Interacting with clients for understanding their requirements
- ✓ Automating Test cases scenarios & Executing using the **VSTS.**
- ✓ Involved in Integration and Execution of Automation scripts.
- ✓ Executed and analyzed the Automation results using **VSTS**
- ✓ Prepared Reviewed the automation scripts.
- ✓ Executing the Test cases and Bug reporting through **MTM and TFS**
- ✓ Follow Agile methodology.