Mohsen **Davarynejad**

CERTIFIED AZURE DATA SCIENTIST · DATA & AI SOLUTION ARCHITECT · PH.D. IN COMPUTATIONAL INTELLIGENCE

Delft, The Netherlands

(+31) 6 593 93 412 | mohsen@thegradient.io | thegradient.io | https://www.linkedin.com/in/davarynejad/ | @mdavarynejad

Summary

Experienced Senior Data Scientist and Consultant with a proven track record in designing and implementing machine learning and AI solutions to solve complex business problems. Proficient in statistical analysis, data visualization, and storytelling, with hands-on expertise in Python (including Numpy, Scipy, Scikit-Learn, and gensim), Matlab, KNIME, PySpark, Microsoft Azure, Databricks, and ML Engineering tools such as Airflow, MLFlow, and Streamlit. Certified Azure Data Engineer (DP-203) and Azure Data Scientist (DP-100). Skilled in delivering ML/AI products and simplifying complex concepts for strategic decision-making.

Skills

Machine learning	Supervised and unsupervised learning, Natural language processing (NLP), Deep learning, NumPy, SciPy, Scikit-Learn, PyTorch, Spark MLlib, SpaCy, Azure ML, AutoML
DevOps	Microsoft Azure, DataBricks, Terraform
Data analysis	Matlab, R, WEKA (an open source collection of Machine Learning algorithms for Data Mining tasks)
Programming	Python, PySpark, C++, LaTeX
Languages	English (Fluent), Persian (Native Speaker), Dutch (Basic)

Career Highlights

- **Full-Cycle Project Leadership and Technical Execution:** Led the end-to-end development and management of several high-impact ML projects, actively participating in hands-on development alongside overseeing all project aspects. Took on roles ranging from project initiation and technical design to implementation while ensuring successful outcomes.
- **Cross-functional Collaboration:** Collaborated with diverse teams including sales, marketing, and technical departments to drive successful project outcomes. Organized workshops with the goal to identify use cases and collaboration possibilities with HR, finance, and facility management departments, fostering a deeper understanding of their needs and enhancing cooperation across departments.
- **Relationship Management:** Managed relationships with multiple vendors and representing my client to ensure the successful delivery of projects. This includes effectively coordinating with vendors to manage their deliverables, ensuring alignment with project timelines and objectives.

Work Experience

Basic Fit International B.V.

Hoofddorp, The Netherlands

SENIOR DATA SCIENTIST - TECHNICAL LEAD

Sep. 2022 - Mar. 2024

Led and contributed to a variety of projects, playing diverse roles ranging from end-to-end project management to solution architecture and team leadership. Highlights of my project experience include:

- Served as Solution Architect for various projects encompassing a range of disciplines such as ”Hierarchical Time Series Analysis” , ”Media Mix Modelling”, and ”Customer Segmentation” to enhance targeting and personalization strategies. This involved making decisions on the choice of platform, components, and algorithms utilized within the projects, and also guaranteeing scalability and the efficiency of the solutions.
- Initiated and managed end-to-end execution of a machine vision project, including project identification, cost estimation, stakeholder management, design, implementation, and delivery.
- Utilized a diverse mix of technologies such as Azure Cloud, Databricks, MLFlow, Real-Time Inferencing, and PyTorch, among others, to drive project success and enhance analytical capabilities.
- Mentored a team of 3 data scientists on advanced statistical techniques, with a focus on text mining methods with application in domains such as Customer Care and HR.
- Provided training in Python and Databricks for enhancing the team’s proficiency in these essential tools for data analysis and processing.
- Facilitated proficiency in various data modalities and problem domains, fostering a collaborative environment for skill development and knowledge sharing within the team and across the proganization.

LinkIT B.V. *De Meern, The Netherlands*

Senior Data Science Consultant - Lead Data Scientist *Jun. 2016 - Aug. 2022*

Worked with major corporations including Royal Dutch Shell, Randstad Groep Nederland, Air France KLM, and Douwe Egberts on a diverse range of projects.

- Led "Basin2Vec" project for data-driven geoscience analogues from text within a 10-week timeframe, employing unsupervised learning models, neural networks, and latent word associations to derive basin similarities. Utilized Python, Genism, Python Dash, Elasticsearch, Azure, PMI, word2vec, doc2vec, clustering techniques such as K-Means and iterative K-Means, dimension reduction methods such as UMAP and PaCMAP to analyze a vast volume of earth science articles for basin analogues.
- Spearheaded "Well surveillance for production optimization" initiative, developing multiple machine learning models to identify early Gas Well production drop signatures and overseeing data engineering components for solution productionization. The project resulted in substantial cost savings and improved operational efficiency. Applied Python, Keras, Python Dash, Databricks, PySpark, and Autoencoders to enhance well surveillance and optimize production processes.
- Focused on "Predictive Maintenance" and "Anomaly Detection", Mohsen actively participated in early anomaly detection of equipment and processes by analyzing sensor data correlations. He played crucial roles in KPI design, model selection, training, and model validation by collaborating with process engineers and business stakeholders. Key technologies and tools used included Python (Numpy, Pandas, Scikit-Learn, Matplotlib), H2O, Keras, Alteryx, Spotfire, Microsoft Azure, Docker, PI data, and Python Dash. The key ML models used in the project were Single class SVM and Autoencoders. A diverse set of data quality checks and data normalization techniques, such as z-score, were applied to ensure the reliability of the analysis.
- Spearheaded the "Predicting LatePax" initiative to address the financial losses resulting from departure delays by determining the probability of late passenger arrivals. Led the identification of relevant datasets, implemented proper data filtering and visualization techniques, and enriched data through feature engineering. Developed and tested multiple machine learning models using Python (Numpy, Pandas, Scipy, Scikit-Learn), Spark, H2O, and R on Hortonworks cluster. Provided insights and recommendations to ultimately minimize the cost of departure delay associated with offloading baggage for late passengers.
- Led a team of three researchers in developing predictive analytics for aircraft arrival times, leveraging expertise in data identification, data augmentation, and machine learning model building. Utilized diverse datasets, including historical and current flight information, geographical, and weather data, to construct and test multiple machine learning models. Achieved a 63% improvement in prediction accuracy compared to scheduled arrival times. Technologies utilized: Hortonworks cluster, Python (Numpy, Pandas, Scipy, Scikit-Learn), Spark, H2O, R, Hive.
- Mohsen was also involved in writing proposals for potential clients and execution of a number of PoCs (e.g: implementing a Chabot using Microsoft bot framework, c# and LUIS.ai). Also he has developed a Twitter bot in Azure called @pyguide.

Capgemini Netherlands B.V. *Utrecht, The Netherlands*

Senior Data Scientist *Aug. 2015 - May. 2016*

Involved in a number of projects including:

- Implementing a recommendation system for an international grocery store on Hadoop Cluster using MLlib on Spark.
- Implementing a recommendation system for an Academy Institution partly on IBM Watson.
- Managing a team of 5 data scientists and developers for an app development project.
- Lean core team member and project manager data science at Philips: Standardizing IT and business processes, VSM, facilitating workshops.

Delft University of Technology *Delft, The Netherlands*

Postdoctoral Fellow in Data Analysis and Data Mining *Aug. 2013 - Jul. 2015*

I was a member of TREsPASS project with the ambition to automate attack tree generation and to prioritise attack scenarios and to suggest the most cost-efficient countermeasures. I was also involved in projects related to the application of data analysis, data mining and machine learning on real-world problems mainly in logistics, traffic control and management.

Daniel den Hoed - Erasmus Medical Center *Rotterdam, The Netherlands*

Postdoctoral Fellow in Mathematical Optimization *Jan. 2013 - Jan. 2015*

Radiation therapy treatment planning is a multi-objective optimization problem. My activity was to refine and extend automated treatment planning with the help of faster optimization algorithms.

- Pareto navigation for multi-criteria IMRT (intensity-modulated radiation therapy) planning.
- Multi-criteria decision analysis for IMRT plan selection.

Technische Universiteit Delft *Delft, The Netherlands*

PhD Researcher in Computational Intelligence *Jan. 2009 - Jun. 2014*

Investigated diverse research topics spanning multiple disciplines and fostering collaboration both within and across departments. Also collaborated with colleagues within the department on projects such as:

- Designed, implemented, and evaluated fast optimization strategies for complex (multi-objective) optimization problems and advanced data analysis algorithms, all implemented in Python, Matlab and R.
- Conducted time series prediction utilizing neural networks and support vector machines, resulting in a minimum 20% enhancement in inflow prediction accuracy for a road traffic network.
- Implemented ensembles of classifiers for supplier segmentation, introducing a more resilient classification approach.
- Analyzed the organizational e-Learning system at Achmea Company, providing insights to aid managers in decision-making processes.
- Applied multi-criteria decision-making methods, including analytic hierarchy process (AHP), across diverse application domains such as measuring landside accessibility of Airports.

Trinité Automation *Uithoorn, The Netherlands*

Software Engineer *Jan. 2009 - Jun. 2009*

Implementation and testing of various distributed road network traffic monitoring and control systems in C++. Results:

- Data collection and aggregation from various sources including Bluetooth detectors, loop detectors, cameras detectors with license plate recognition.
- Estimation of missing data using Treiber-Helbing filter. The fitter has been used to have a better understanding of the current state of the network.
- Implementation and testing of FileProof project (for Traffic Control Center of the Province of North-Holland - the project includes 2462 sensors and 1599 actuators which are distributed in the network of 80 by 50 km).

Tous Stadt Co. Consulting Engineers *Mashhad, Iran*

Software Engineer *Jun. 2008 - Jan. 2009*

Implementation and testing of a Short Circuit Analysis software for Power Grid (substations) in Matlab. Results: Automation in analysing the effect of 3-phase, 1-phase, line-to-line and line-to-ground fault currents on distribution substations.

Education

PhD in Artificial Intelligence *Delft, The Netherlands*

Delft University of Technology *2009 - 2014*

Thesis: Deploying metaheuristics for global optimization

Master of Science in Control Systems Engineering, Summa cum laude *Mashhad, Iran*

Ferdowsi University of Mashhad *2004 - 2006*

Thesis: Fuzzy Fitness Granulation in Evolutionary Algorithms for Complex Optimization

Bachelor of Science in Control Systems Engineering *Mashhad, Iran*

Ferdowsi University of Mashhad *1998 - 2003*

Thesis: Intelligent Greenhouse, design and implementation

Certificates

Sep. 2021 **Azure Data Engineer (DP-203) - Associate**, Microsoft *991874506*

Aug. 2021 **Azure Data Scientist (DP-100) - Associate**, Microsoft *991874506*

Jul. 2024 **Databricks Certified Developer for Apache Spark – Associate**, Databricks - Planned for July 2024

ADDITIONAL INFORMATION

Supplemental Education

Attendee, TRAIL Research School on Transport, Infrastructure and Logistics on:

- Data-analysis and statistics
- Logistics and freight transport systems analysis
- Advanced inventory theory

Publications (co-)author of over 35 refereed publications on stochastic search methods, traffic network prediction and control. The complete list of publications is available on my personal website.

Projects Involved in a number of European projects including Con4Coord and TREsPASS.

Thesis Supervision During my PhD research at TU Delft, I supervised 5 Master research projects in various domains including: Mining techniques in cyber-attacks, Measuring operational performance of airport traffic and transportation, Exchange rate prediction, and Inventory management.

Hobbies Hiking, Basketball, Playing Chess and Twitting/Blogging. My most recent hobby is to maintain and improve my own twitter bot @pyguide.