Saleh **Mozafari**

Data Scientist | Machine Learning Engineer

7 Aug 1981 – Tehran, Iran linkedin.com/in/salehmozafari
+49 152 26988580 mozafari.dev@gmail.com
Bännjerstraße 73, 67655 Kaiserslautern, Germany

With over a decade of experience as a Data Scientist, I am not only proficient in AI, Machine Learning, and Computer Vision but also a skilled full-stack developer, capable of building end-to-end solutions. My Master's in Artificial Intelligence and expertise in Python, C++, and advanced analytics have enabled me to develop predictive models and leverage machine learning frameworks effectively. I excel in visualizing complex data insights and have a strong foundation in cloud technologies. My problem-solving abilities are matched by my capacity for cross-functional teamwork and exceptional communication skills, making me a versatile asset in digital transformation projects. My dual competencies in both data science and full-stack development, combined with a strategic approach to problem-solving and a knack for clear, impactful communication, position me uniquely to drive operational efficiency and innovation within dynamic industry environments.

Work Experience

Present
Sep 2018

Expert Data Scientist | Machine Learning Engineer – **WidasConcepts GmbH**, Wimsheim, Germany

Project: Cidaas-ID-Validator

- Directed the development and deployment of scalable microservices to operationalize machine learning models, enhancing system efficiency and responsiveness.
- Innovated an OCR (Optical Character Recognition) model tailored for high-precision information extraction from various identity documents, significantly improving the accuracy of data retrieval processes.
- Engineered a sophisticated security check framework for identity verification, integrating Computer Vision and Probabilistic Graphical Models to detect fraudulent documents effectively.
- Pioneered the design and development of a Human-Computer Interaction (HCI) application leveraging eye-tracking methodologies (ninoxipy) for real-time liveness detection, setting new standards in identity verification technology.
- Played a pivotal role in the conceptualization and implementation of an MLOps framework, facilitating seamless model integration, monitoring, and management, thereby ensuring continuous improvement and operational excellence of deployed machine learning solutions.

Docker | Kubernetes | GitLab | DC/OS | Python | C++ | R | OpenCV | dlib | Ray | tesseract | ocrupy | Flask | Jupyter-notebook | numpy | scipy | pandas | matplotlib | seaborn | PyTorch | Pyro | Tensorflow | keras | scikit-learn | scikit-image | scikt-fuzzy

Present
2020

Lead Data Scientist | Machine Learning Engineer – **WidasConcepts GmbH**, Wimsheim, Germany

Project: Cidaas- FDS (Fraud Detection System)

- Architected an ML-based cybersecurity solution, focusing on the integration of advanced fraud detection mechanisms within the system's core to bolster security measures.
- Engineered RESTful APIs to seamlessly serve fraud detection models, enabling real-time fraud analysis and immediate response to security threats.
- Instrumental in the development and maintenance of a CI/CD pipeline, ensuring the agile deployment and rigorous testing of models, which contributed to a streamlined development process.
- Innovated a Smart Multi-Factor Authentication (MFA) system utilizing a Fuzzy Inference System (FIS), which enhanced user verification processes by adapting to varying risk levels.
- Conducted Causal Inference analysis to identify the root causes of fraudulent activities, enabling the implementation of targeted preventive measures.
- Achieved a breakthrough in fraud detection accuracy, reaching a very high accuracy rate of 97%, significantly reducing the risk and impact of fraudulent activities on the platform.

MLFlow | scikit-learn | scikit-image | scikt-fuzzy | CausalInference | Docker | Kubernetes | GitLab | Python | R | FastAPI | Jupyter-notebook | numpy | scipy | pandas | matplotlib | seaborn | MongoDB | Elasticsearch | Kibana

2022
2020

Lead Data Scientist | Machine Learning Engineer – **WidasConcepts GmbH**, Wimsheim, Germany

Project: Bosch-EBR (Experience-based Repair)

- Pioneered the development of an NLP-based vehicle identification system, employing Named Entity Recognition (NER) and Graph Theory to accurately identify and classify vehicle components.
- Architected and implemented a RESTful API, facilitating the efficient delivery of the vehicle identification model to stakeholders, enhancing user accessibility and system integration.
- Designed and developed a robust stream processing framework using Faust for EBR microservices, optimizing data handling processes and improving the system's scalability and resilience.
- Contributed the creation of an LLM-based Concept Identification Service, leveraging large language models to mine car part failure discussions in forums, and algorithmically identifying approved solutions to address these failures. This innovative service significantly improved the efficiency of diagnosing and resolving vehicle issues by directly linking reported problems with validated solutions.

Python | NLTK | SpaCy | Gensim | Faust | FastAPI | Elasticsearch | Kibana

2019
2018

Data Scientist – **WidasConcepts GmbH**, Wimsheim, Germany

Project: Porsche-CCD (Corner Case Detection)

- Engineered a state-of-the-art spatio-temporal model that leverages internal sensory data and the geolocation of Porsche cars to detect corner cases, significantly enhancing the predictive capabilities and safety measures.
- Spearheaded the design and development of an interactive UI for monitoring Porsche cars, incorporating ReactJS to create a responsive and user-friendly interface that enables real-time visualization of data and identification of plausible corner cases.
- Developed a high-performance data streaming system utilizing MQTT and Tornado, ensuring efficient data transmission and processing capabilities for real-time monitoring and analysis.
- Implemented advanced data visualization techniques using Plotly and Bokeh, enabling complex data insights to be comprehensively understood through interactive graphs and charts, thus facilitating informed decision-making processes.

Python | ReactJS | Bokeh | Plotly | Tornado | MQTT | scikit-learn

2019
2018

Computer Vision Engineer | Data Scientist – **WidasConcepts GmbH**, Wimsheim, Germany

Project: Chinese Painting Seal Assessment

- Pioneered the development of a cutting-edge deep learning model using state-of-the-art Convolutional Neural Networks (CNNs) in TensorFlow to accurately identify seals of Chinese artists, enabling the authentication of artworks with high precision.
- Innovated a recurrent neural network model designed to assess the originality of artists' seals, effectively distinguishing genuine artifacts from forgeries. This model leveraged the nuanced patterns and historical data of seal imprints, contributing significantly to the preservation of cultural heritage and the art market's integrity.

Python | Tensorflow | Keras | RNN | CNN | OpenCV | scikit-image | scikit-learn | JupyteLab | Pillow

Aug 2018
Sep 2014

Researcher and Developer – **DFKI GmbH**, Kaiserslautern, Germany

- Spearheaded the innovative modeling of gaze movements within the Immersive Quantified Learning lab (iQl), advancing the field of Human-Document Interaction through groundbreaking research that bridges the gap between digital learning environments and natural user engagement.
- Led the design and implementation of a cutting-edge Human-Computer Interaction system utilizing eye-tracking technology for the AICASys Project, markedly enhancing the accessibility of digital interfaces for individuals with disabilities, thereby democratizing technology use.
- Developed a novel generative probabilistic model for synthesizing eye-movement patterns, which stands as a significant contribution to understanding and predicting user interactions in digital environments, paving the way for more intuitive and adaptive systems.
- Contributed to academia as a Teaching Assistant for the Data Mining Course at TU Kaiserslautern, enriching the educational experience of students with practical insights and fostering a deep understanding of critical data analysis techniques.
- Supervised Master and Bachelor dissertations in the fields of AI and HCI at TU Kaiserslautern, guiding students through their academic journeys with mentorship that emphasizes innovation, critical thinking, and real-world application of theoretical knowledge.

Python | C++ | R | Matlab | ZMQ | IoT | OpenCV | dlib | jupyter-notebook | numpy | scipy | pandas | matplotlib | seaborn | pillow | Edward | keras | scikit-learn | scikit-image | scikt-fuzzy | Git | TortoiseSVN

Aug 2014 – Nov 2013 | **Full-stack Developer** – **Yareegar Clinic**, Tehran, Iran

- Designed a comprehensive personality assessment project for Mapna Group, successfully developing and deploying a robust evaluation system for over 20,000 employees. This initiative significantly contributed to enhancing workforce management and organizational development through tailored psychological insights.
- Demonstrated exceptional full-stack development skills by designing and developing a dynamic portal for psychological assessments. Utilized PHP for server-side logic and Python for data analysis and processing, showcasing a versatile proficiency in both front-end and back-end technologies. This portal facilitated seamless, efficient assessments, highlighting my ability to deliver comprehensive digital solutions from concept to deployment.

Python2 | PHP | XAMPP | JavaScript | CSS | Autobahn|Python | TortoiseSVN | MySQL

Nov 2013 – Sep 2012 | **University Lecturer** – **Azad University**, Tehran (Roudehen), Iran

- As a dedicated university lecturer, I have passionately imparted knowledge in cutting-edge subjects such as Machine Learning, Data Engineering, and Computer Vision to undergraduate Computer Science students. My approach to teaching these courses emphasizes not just the theoretical underpinnings but also practical applications, preparing students to tackle real-world challenges with innovative solutions.
- For postgraduate Clinical Psychology students, I have taught Data Analysis and Advanced Statistics, focusing on equipping them with the quantitative tools necessary to conduct rigorous research and analysis in their field. My methodology integrates statistical theories with hands-on data analysis, fostering a deep understanding of complex concepts and their application in clinical settings.

Python | R | SPSS | Matlab

Ausbildung

2009 – 2011 | **M.Tech in Computer Kognition Technologie** - University of Mysore, India
Titel der Dissertation: Cancer Prediction Using Microarray Expression Data
Notendurchschnitt: 9.3 / 10 *(Ausgezeichneter)*

2000 – 2004 | **B. Sc in Software Engineering** - Azad-Universität, Iran
Graduation project: E-learning platform mit ASP.Net
Notendurchschnitt: 15.13 / 20 *(Iranische Leistungsbeurteilung)*

Skills

Programming Languages	**Python**, R, C, C++, JavaScript, Go
Libraries	**Numpy**, Scipy, **Pandas**, **scikit-learn**, scikit-fuzzy, scikit-image, **OpenCV**, dlib, **PyTorch**, Causal-Inference, Tensorflow, Keras, Statsmodels, Django, Flask, FastAPI, Sphinx
NLP and LLM	NLTK, SpaCy, Gensim
Visualization Tools	**matplotlib**, **seaborn**, bokeh, plotly, ggplot2, Tableau
Databases	MySQL, PostgreSQL, MongoDB, Redis, Elasticsearch
IDEs	**PyCharm**, **Visual Studio Code**, Sublime-text, RStudio
MLOps Tools	**GitLab CI/CD**, **Docker**, Kubernetes, MLFlow, Apache Kafka
Operating System	Fedora, **Ubuntu**, MacOS, Windows

Honors and Awards

- Distinction student in Master of Computer Technology with GPA 9.3/10, University of Mysore, India.
- Best paper award in IAPR-HDI conference, Kyoto, Japan, Nov 2017.

Languages

English	Business fluent
German	Limited business fluent (B1)
Persian	Native

Publications

1. **Saleh Mozafari**, Pascal Klein, Mohammad Al-Naser, Stefan Küchemann, Jochen Kuhn1, Thomas Widmann, and Andreas Dengel. Quantifying Gaze-based Strategic Patterns in Physics Vector Field Divergence. In Agents and Artificial Intelligence. Lecture Lecture Notes in Artificial Intelligence book sub series (LNAI). Springer International Publishing, 2021. (Under publication)
2. Pascal Klein, Jouni Viiri, **Saleh Mozafari**, Andreas Dengel, and Jochen Kuhn. Instruction-based clinical eye-tracking study on the visual interpretation of divergence: How do students look at vector field plots? In Physical Review Physics Education Research 14 (1), 010116, American Physical Society,2018.
3. **Saleh Mozafari**, Pascal Klein, Mohammad Al-Naser, Stefan Küchemann, Jochen Kuhn1, Thomas Widmann, and Andreas Dengel. Classification of Visual Strategies in Physics Vector Field Problem-solving. In Proceedings of the 12th International Conference on Agents and Artificial Intelligence - Volume 2: ICAART, 2020.
4. **Saleh Mozafari**, Pascal Klein, Jouni Viiri, Sheraz Ahmed, Jochen Kuhn, and Andreas Dengel. Evaluating similarity measures for gaze patterns in the context of representational competence in physics education. Proceedings of the 2018 ACM Symposium on Eye-tracking Research and Applications (ETRA), 2018.
5. **Saleh Mozafari**, Federico Raue, Saeid Dashti Hassanzadeh, Stefan Agne, Syed Saqib Bukhari, Andreas Dengel. Reading type classification based on generative models and bidirectional long short-term memory. International Conference on Intelligent User Interface (IUI), UISTDA Workshop, Japan, Tokyo, 2018.
6. Marc Beck, **Saleh Mozafari**, Syed Saqib Bukhari, Andreas Dengel. Landscape or Portrait? The Impact of Page Orientation on the Understandability of Scientific Posters. 14th IAPR International Conference on Document Analysis and Recognition (ICDAR), 2017.
7. **Saleh Mozafari**, Mohammad Al-Naser, Syed Saqib Bukhari, Damian Borth, Shanley EM Alleny, Andreas Dengel. An eye movement study on scientific papers using wearable eye-tracking technology. 2016 Ninth International Conference on Mobile Computing and Ubiquitous Networking (ICMU), Kaiserslautern, Germany, 2016.
8. **Saleh Mozafari**, Pascal Klein, Syed Saqib Bukhari, Jochen Kuhn, Andreas Dengel. Entropy-based transition analysis of eye movement on physics representational competence. Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing, WAHM, 2016
9. **Saleh Mozafari**, Pascal Klein, Syed Saqib Bukhari, Jochen Kuhn, Andreas Dengel. A study on representational competence in physics using mobile eye-tracking systems. Proceedings of the 18th International Conference on Human-Computer Interaction with Mobile Devices and Services, 2016.
10. **Saleh Mozafari**, Syed Saqib Bukhari, Andreas Dengel. Analysis of Text Layout Quality Using Wearable Eye-trackers. IEEE International Conference on Multimedia and Expo Workshops (ICMEW), 2015.
11. Mohammad Al-Naser, Peter Lanzer, Andreas Dengel, Syed Saqib Bukhari, **Saleh Mozafari**. Knowledge transfer from experts to novices in minimally invasive catheter-mediated (MIC) interventions, eye-tracking study. Proceedings of the 18th International Conference on Human-Computer Interaction with Mobile Devices and Services, 2016.
12. Mohammad Al-Naser, **Saleh Mozafari**, Syed Saqib Bukhari, Damian Borth, Andreas Dengel. What Makes a Beautiful Landscape Beautiful: Adjective Noun Pairs Attention by Eye-Tracking and Gaze Analysis. Proceedings of the 1st International Workshop on Affect and Sentiment in Multimedia, 2015.
13. Marco Stricker, Syed Saqib Bukhari, Mohammad Al Naser, **Saleh Mozafari**, Damian Borth, Andreas Dengel. Which Saliency Detection Method is the Best to Estimate the Human Attention for Adjective Noun Concepts? International Conference on Agents and Artificial Intelligence (ICAART), 2017.
14. M Mohammadi, **Saleh Mozafari**, Aradhya VN Manjunath, G Hemantha Kumar. An Improved Handwritten Text Line Segmentation Technique. Advances in Computing and Communications. Pages 289-296, 2011.