AHMET F CAKMAK, PH.D.

Senior Data Scientist and Senior Software Engineer
PHONE: +1(470)-695-5521   EMAIL: cakmakahmetf@gmail.com

EDUCATION:

- Yıldız Technical University, Istanbul, Turkey, 09/09 – 02/14
  - Ph.D. in Mathematical Engineering
- Gebze Technical University, Kocaeli, Turkey, 09/05 – 11/07
  - M.Sc. of Mathematics
- Uludag University, Bursa, Turkey, 09/00 – 06/04
  - BS of Mathematics

SKILLS:

Programming

Python, R, C#, .NET Core, C++, Pascal, LATEX, SAS, MATLAB, HTML, JavaScript.

AI / Machine Learning / NLP

NLP, Information Retrieval, NLTK, word2vector, Regression, PCA, Signal Processing, Decision Trees, Clustering, Deep Learning, PyTorch, TensorFlow, Keras, CNN, sci-kit-learn, Pandas, NumPy, LibSVM, JupyterLab, GitLab.

Software

MySQL Server, Power BI, .NET Framework, PostgreSQL, SQL, Docker Kubernetes, Azure Microservices, Apache Kafka, GIS, API, Apache Hadoop, Apache Spark, Amazon Web Services (AWS), Google Cloud, Git, and GitHub.

EXPERIENCE

07/2023 - Now
Motorola Solutions, Atlanta, GA (Remote)
Senior Data Scientist

- Contributing both mathematical and programming support for advanced NLP challenges for live stream and post-processed incident data.
- Building sentiment analysis apps and toolkits with NLTK-type libraries and Large Language Models and productize the models regarding customer and corporate purposes.
- Analyzing and creating Monte - Carlo type simulations and probablistics approaches that resulted from the concerning ML - NLP problems by establishing them from theory to application using MATLAB, R, Python and other programming languages.

03/21 - 07/2023
Motorola Solutions, Atlanta, GA (Remote)
Senior Software Engineer

- Creating enterprise-grade Client/Server applications and services in Azure DevOps following an Agile methodology in Python, .NET FW, C#, Javascript, and SQL, and writing efficient and SOLID codes in those languages.
- Implemented scalable microservices using .Net Core, Azure Functions, Azure Service Azure Auto ML, Docker Containers, and Kubernetes. Maintained build and release pipelines on Azure DevOps.

- Building web applications of NLP models, supported by neural networks and transformers, for text/audio data parsing to report the trends and, building data-driven web applications with REST APIs, Fast API and Flask by using and mastering ML and NLP models.
- Hosting ML and NLP applications on Azure Databricks cloud platform by supporting GitHub integration followed by End to End MLOps operation with MLFlow, Feature Store, and Auto ML.
- Create and execute unit, integration, system, load, and acceptance test plans and scripts. Use software system testing procedures and microservices tools to upload the obtained results for production. Use Databricks and Azure DevOps and CI/CD integration.
- Perform modeling, designing, and coding activities, employing structured methods. Prepare design documentation for data science and analytics of the software development process.

01/22 - 06/22
Carnegie Mellon University / Emeritus, Cumming, GA
Subject Matter Expert for NLP

- Selected as an instructional member of Carnegie Mellon University School of Computer Science Executive Education's Introduction to Natural Language Processing program that requires part-time contribution.

- Create tutorials and find/create supplemental readings. Review faculty-filmed videos. Help writing the connective text for the Canvas instance of the course. Create various types of assessments for the system. Attend a weekly meeting with the design team. Build end-to-end NLP hands-on programming projects for each chapter that includes deep learning and transformers, and test the python programming activities by unit testing.

12/18 – 03/21
Vanderbilt University Medical Center, Nashville, TN
Staff Scientist
- Create a classifier that distinguishes healthy from unhealthy infected B-cell repertoires. Predict liabilities in antibody sequences. Data mining antibody sequences from the publicly available patent database.

- Constitute NLP models for semantic knowledge representation, knowledge mapping, and sentiment analysis for predictive analytics. Apply deep learning tools such as TensorFlow, Keras, and PyTorch and implement deep natural language processing results by word2vector, sentiment analysis, machine translation, and Seq2Seq. Phenotyping biomedical data from semantic knowledge representation with GWAS (Genome-Wide Association Studies) and PheWas (Phenotype Wide Association Studies) for machine/statistical learning results.
- Lead development of geocoding, and geospatial data analysis program QGIS by managing docker container coding files. Data mining by MySQL/PostgreSQL for project managers. Creating clustering classifications for the providers applying, sci-kit learn, KNN, Gaussian MM, and Random Forest algorithms.
- Build AI algorithms and cloud solutions to utilize precision medicine algorithms in the All of Us Program over the Google Cloud Platform and provide coding support to implement complex problem solutions.

08/16 – 12/ 18
Middle Tennessee State University, Murfreesboro, TN
Lecturer Professor
- Lecturing for calculus, applied statistics, and various mathematics courses using active learning to visualize real-life problems by formulas.
- Constructing an optimized Q-Learning driving agent that navigates a smart cab toward the destination.
- Evaluate the performance and predictive power of a model on data collected from homes in Boston, MA.
- Applied unsupervised learning techniques on product spending data collected for customers of a wholesale distributor to identify customer segments hidden in the data.

10/15 – 08/16
Tennessee State University, Nashville, TN
Researcher Scientist
- Developed theory and algorithms addressing subspace segmentation projects sponsored by the US Army.
- Maintain a framework for finding similarity matrices for the segmentation of data that can be used in clustering algorithms.
- Analyze graph connectivity of data nodes to develop the theory for a general structure via SVD and Low-Rank Representation (LRR).
- Provide advanced mathematical and coding support to experiment outcomes and present the findings of the research project at international conferences.

09/14 – 10/15
Vanderbilt University, Nashville, TN
Researcher Scientist
- Engaged in research that will be resulting in a solution to Dynamical Sampling problems which is applicable in Wireless Sensors Networks (WSN) for signal reconstruction.
- We have developed a model for normal operators in Hilbert space to use various applied and computational Harmonic Analysis techniques.

01/09 – 09/14
Yildiz Technical University, Istanbul, Turkey
Research & Teaching Assistant
- Research in Applied Math and Computer Science, application of Statistical/Machine Learning. Attend and present advanced research developments and publications at international scientific symposiums.
- TA for undergraduate classes on Calculus, Statistics, Programming Languages (on C++, Python, MATLAB, SQL), Object-Oriented Programming, Differential Equations, Algorithms, Data Structures, and Optimization Techniques.
- Lead and validate student internships on Applied Software/Computer Engineering-Based Internship with industry-related firms.

PROJECTS

03/21 - Now
Motorola Solutions, Atlanta, GA (Remote)
- Unit testing with Navigator API system to measure and verify transcription and document retrieval services.

- Integrating shapefiles with Power BI map features including ArcGIS, Map Box, and other visual tools.

- Web application building and hosting them in Azure Databricks by maintaining the whole infrastructure.

- Creating data-driven NLP apps with APIs based on classification by using open-source transformers and various NLP libraries.

12/18 – 02/21
VUMC Biomedical Projects, Nashville, TN
- Using quantitative models of the stochastic V(D)J recombination process that have been inferred from repertoire data. Apply machine learning models using various features were constructed as well as a gradient boosting machine (GBM) classifier combining models for DNA/RNA sequencing.

- Apply NLP models for Semantic Knowledge, Representation, Knowledge Mapping, and Sentiment Analysis for adverse drug effect analysis. Phenotyping UMLS ids for the PheWas algorithm to obtain statistically significant signals from the flow of the corpus.
- Build NLP Algorithms and cloud solutions, to utilize precision medicine algorithms in the All of Us Program over the Google Cloud Platform and provide coding support for the group.
- Contribute hands-on experience in precision medicine of geocoding, QGIS, and various clustering algorithms. Design, build and implement prediction, clustering, and optimization algorithms to implement healthcare/business problems.

07/17 – 06/18
UDACITY Machine Learning Program, Nashville TN
- Built a deep learning pipeline that can be used within a web or mobile app to identify real-world, user-supplied pictures for image recognition that obtained 87% accuracy.
- Construct a Support Vector Machine (SVM) model for gathering randomly collected audio files to be able to extract digits out of each conservation. According to our tests, we reached 98% accuracy for individual digit recognition precision and around 89% accuracy for entire digit recognition.

10/15 – 08/16
Subspace Segmentation, Nashville TN
- Maintain a framework for finding similarity matrices for the segmentation of datasets that can be used in clustering algorithms.
- Analyze graph connectivity of data nodes to develop the theory of a general structure via Singular Value Decomposition (SVD) and Low-Rank Representation (LRR).

PORTFOLIO:

https://github.com/cakmakaf, https://www.linkedin.com/in/ahmet-faruk-cakmak/, https://scholar.google.com/citations?user=AxMLsxQAAAAJ&hl=en