https://gmihaila.github.io
linkedin.com/in/mihailageorge
https://gmihaila.medium.com

George Mihaila
Data Science | Machine Learning |Software Engineer

georgemihaila@my.unt.edu
(302) 864-7518

PROFESSIONAL EXPERIENCE

TCF Bank – Huntington Bank

Data Scientist, Part-Time *Jan 2021 – Present*

- Translated credit risk models into python and R with unit testing implementation.
- Built a Proof of Concept (PoC) application around Machine Learning (ML) Operations (MLOps) framework on Kubernetes: automated model registry and deployment using MLFlow, data pipelines with Airflow, API model calls using FastApi and CI/CD pipeline in Jenikins.
- Helped transition from MLOps solution vendor that reduced model deployment cost to zero.

State Farm

Data Scientist, Intern *May 2020 – Jul 2020*

- Created a PoC application for stock portfolio optimization using Reinforcement Learning (RL) in Python.
- Built a RL framework tool that makes it transferable in different business areas to help create RL environments and models.
- Helped create a PoC application that optimizes claims automatic payment process using RL in Python.

Machine Learning Engineer, Intern *May 2019 – Jul 2019*

- Automated modeling evaluation in python: build baselines and compare against target model reducing deployment timeline by 90%.
- Wrote and published a python package for automatic model evaluation that can be used across different departments.
- Built python program to generate HTML reports on baseline model evaluation and data analysis plots integrated in a CI/CD pipeline.

Data Scientist, Intern *May 2018 – Aug 2018*

- Used Natural Language Processing (NLP) tools in python to extract meaningful features from claims text data that were used in addition to existing claims feature to increase model baseline performance in predicting claims duration.

University of North Texas

Data Scientist *Sep 2018 – May 2020*

- Helped research labs with model implementations, debugging and attract more funding.
- Helped maintain latest ML and deep learning frameworks for University of North Texas High Performance Computing (HPC) center.
- Created and held workshops and tutorials on NLP and using HPC services that helped increased userbase by 20%.

Teaching Assistant *Sep 2017 – May 2018, Sep 2020 – May 2021*

- Building and debugging C/C++ coding assignments and help coordinate coding exams.
- Peer mentor and grader for students in Computer Science basic and advance courses.

Machine Learning Researcher *Jan 2017 – Present*

- Implement concepts in python from research papers related to NLP and Computer Vision (CV).
- Building state of the art deep learning models in python in the domains of language modeling and image classification.

EDUCATION

University of North Texas

- **Doctor of Philosophy (PhD) in Computer Science** *Dec 2021*
  Research Areas: ML, deep learning, NLP, CV.
- **Master of Computer Science** *Dec 2019*
  Relevant Coursework: Machine Learning, Deep Learning, Big Data, Natural Language Processing, Image Processing.

SKILLS

Proficient	Need Warmup	Intermediate
Python, Tensorflow, PyTorch	Java, C, C++, Matlab, R	SQL, Docker, AWS, Swift, Android, Hadoop, HTML

PERSONAL PROJECTS

- Held a webinar titled *Text Classification using GPT-2 and PyTorch* hosted by AICamp with over 300 participants.
- Technical reviewer for one of the first transformers models book titled *Transformers for Natural Language Processing in Python*.
- Contributed to open-source software on GitHub. Most important was Hugging Face *Transformers* and *Datasets* libraries.
- Created tutorials on state-of-the-art NLP transformers models like Bert and GPT-2. Find them at https://github.com/gmihaila/ml_things.