Muhammad Khalifa

33, Fostat St, Cairo, Egypt • (+20) 114 522 3109
Muhammad.e.khalifa@gmail.com

Employment

Machine Learning Engineer **Sypron Solutions** **04/2018 – Present**

Research and Development of Anomaly Detection models for Predictive Maintenance.
Built and tested multiple architectures for time-series anomaly detection including LSTM classifier and autoencoder using *TensorFlow*.
Deployed real-time anomaly detection models with *TensorFlow-Serving API* and *Flask*.
Implemented stream processing and analytics of the incoming sensory data stream with *Docker*, *VerneMQ*, *Kafka-MQTT Connect*, *Kafka* and *Spark Streaming*.

Software Engineer (Full Stack) **CITC Mansoura University** **7/2017 - 10/2017**

Worked as part of a team in CITC on Equipment's Maintenance System for Oncology Center of Mansoura University.

Implemented main functions of the systems such as equipment's archiving and placing (HTML, CSS, Node JS and MySQL).

Teaching Assistant **Mansoura University** **1/2017 – 5/2017**

Courses: Introduction to Computer Programming (Visual Basic) for freshman students.

Handled guiding and evaluating students' assignments. Assisted with final exams grading.

Education

Mansoura, Egypt **Mansoura University** **Fall 2011 – June 2016**

B.Sc. in Computer Systems Engineering.
Grade: 87.0%. (Distinct with Honors).
Rank: 8 / 119.

Cairo, Egypt **Cairo University** **Spring 2018 – Now**

Full-time MSc. in Computer Science.
Finished one year of masters-level studies. Excellent (90.1%) average grade in all courses.
Thesis Topic: *Transfer Learning for Arabic Natural Language Processing in Low-Resource Scenarios.*

Research

- Primary research interests are: Machine Learning, Natural Language Processing.
- Character Convolutions for Arabic Named Entity Recognition with LSTM Networks with professor Khaled Shaalan (British University of Dubai). Published at *The Journal of Computer Speech and Language.*
- Book Success Prediction with CNNs and Readability Scores. With professor *Aminul Islam* (University of Louisiana at Lafayette). Submitted to *EMNLP-2019.*

LANGUAGES AND TECHNOLOGIES

- Python; C++; C; Java; SQL; JavaScript; MATLAB.
- TensorFlow; Keras; PyTorch.
- Flask; CSS; Bootstrap; NodeJS.

PROJECTS

- Implemented ULMFit training and fine-tuning scheme with PyTorch including Discriminative Fine-tuning and Slanted Learning rates. Trained on Arabic Wikipedia, fine-tuned on Arabic Dialect Identification AOC dataset. Got an accuracy of 80% on word-level LM and 82% on character-level LM (4-5% improvement over not using Transfer Learning). Link.

- Bilateral Multi Perspective Matching for Natural Language Sentences. Implemented the paper model with PyTorch. Got a validation accuracy of 85% on Quora Question Pairs Dataset. *Link*

- Arabic to English machine translation with the Transformer model (6 self-attention blocks, 8 attention heads). Dataset used: OpenSubtitles v2018 (~262M tokens). Currently training and evaluating the model. *Link*

- Siamese CNNs for Duplicate Question Detection on Quora Question Pairs Dataset. Got an accuracy of 82% on a validation set.

- Arabic News Summarization with Seq2Seq model using PyTorch. Dataset used was SaudiNewsNet. Got good summaries but overall bad BLEU score due to the small size of the dataset. *Link*

- Movie Genre Prediction: Predicting movie genre from title using Naive Bayes and Support Vector Machines with bag of words model using word2vec. Got an f1-score of 63%. *Blog*

- Arabic Poetry Generation: Poetry generation model based on a word-level language model with TensorFlow using a 2-layer LSTM network. Corpus used: 10000 Arabic dialect poems scraped from *www.adab.com*.

- Smart Exam Generator: An Application for teachers to generate exams based on their preferences. Used Genetic Algorithms to optimize generated exams. Used a combination of Flask, JavaScript and CSS.

- Traffic Count Prediction with 1-hour intervals with GMLE (Gaussian Maximum Likelihood Estimation). Dataset used: Seattle Traffic Flow Counts.

- University Timetable Scheduling with Genetic Algorithms. Timetables are generated to optimize a certain objective function that prohibits infeasible schedules.

ADDITIONAL EXPERIENCE AND AWARDS

- Won the first place in the IDAT@Fire 2019 competition (as YOLO team) for Irony Detection in Arabic Tweets among 18 competing teams.
- Won the first place in IEEE Code Door 2014 algorithmic problem-solving competition.
- Won 1st place Nasa's Space Apps Challenge 2016 edition in the Journey to Mars Challenge.