KHEM SOK

1420 Tyson Ave, Philadelphia, PA 19111 | 215-460-7426 | khemsok97@gmail.com | https://khems.io/ | LinkedIn

---

EDUCATION

University of Michigan, Ann Arbor, MI   Aug 2020 – Present
Master of Science, Data Science
Pennsylvania State University, State College, PA   Aug 2014 – May 2018
Bachelor of Science, Computer Engineering

TECHNICAL SKILLS

Programming Languages: **Python, JavaScript, TypeScript**, Java, R, HTML5, CSS3, C++, C
Technologies: **React, React Native, Redux, NodeJS, GraphQL**, PySpark, Docker, PyTorch, NiFi, PySpark, Flask, Nginx, Tableau
Cloud Computing: Azure (Function, ML Service, Key Vault, Databricks, Data Lake, Data Factory, etc.) , AWS (EC2, Lambda, RDS)
Database: Microsoft SQL Server, Postgres MySQL, Firebase, MongoDB, Mark Logic

---

WORK EXPERIENCE

❖ **Software Engineer II**, *DuPont*   April 2020 – Present
- Build a web application that serves as an **analytics tool for business** to search for documents, display complex interactive visualizations such as market trends, heatmaps, etc. and get the latest news data from SharePoint using React, Redux, NodeJS, MarkLogic, and NiFi.
- Design and build a productionized machine learning pipeline for forecast models via **Databricks orchestrated by Azure Data Factory** with data coming from SAP storing in **Azure Data Lake** to be transform and clean into **Delta Lake** format for structured consumption and automatically upload forecast results to Sharepoint using Microsoft Graph API.
- Lead a team to build an ETL pipeline to ingest data from an endpoint and enrich it user information coming from Azure Active Directory via Microsoft Graph API and store it on Azure SQL to be consumed by Power BI.

❖ **Software Engineer I**, *DuPont*   Jan 2019 – April 2020
- Create, design, and **implement data science and data engineering pipelines** on manufacturing minute and event-based level data from IP21 historian to ingest, merge, transform, clean, perform feature engineering, model implementation, model evaluation, and model reports. Built using a pipeline tool called Kedro.
- Design and implement a web-based application that allows operators to use machine learning models to get recommendations on a manufacturing process. Instantly see **increase in production by 15%** with one week of usage. Built using Flask and Dash with data coming from SQL Server and IP21 historian.
- Create, design, and implement an **ETL pipeline for data processing and transformation** from machines' raw CSV files to Azure SQL database using Azure Microservices including Azure Function, Azure Storage, and Azure Key Vault
- Apply natural language processing algorithm **to extract keywords and sentiments** from company's annual surveys to better understand employees' takes and **produce visualization to easily capture the emotions with Azure Text Analytics**.
- Create **and design a chatbot** that is able to intelligently communicate with users' request built upon **Azure Bot Framework with LUIS and QnA Maker** in NodeJS and React

❖ **Teacher Assistant** (C++/Python), *Pennsylvania State University*   Sept 2015 – Aug 2018
- Assists students with C++ and Python related coursework, formulate end of the year class related projects, grade assignments

---

PERSONAL PROJECTS

❖ **YourBooks (Link) (Web GitHub) (Mobile GitHub) (Google Playstore)**
Application that allows user to discover and **gets recommended new books**, track their books list, take notes, and more. ML model deployed in **AWS Lambda**, Firestore as the database. Web App built with NextJS. Mobile application runs **on both iOS and Android** built with React Native. Wrote script to capture 5,000+ Google books data (descriptions, genre, authors, etc.) to train the recommender model.

❖ **Facial Expression Detector (Medium) (GitHub)**
Design a machine learning model that is able to detect 4 facial expressions with **transfer learning** built with PyTorch.

❖ **RandomThoughts (Link) (GitHub)**
Web application that allows users to start writing down and storing thoughts using Firebase as the DB and Google OAuth as the authentication.

❖ **Deploy ReactJS and NodeJS in AWS EC2 with Docker and Nginx (Medium) (GitHub)**
Guide on how to deploy ReactJS and NodeJS application in AWS EC2 with Docker and Nginx.

❖ **Redux Is Simple (Link) (Medium) (GitHub)**
Guide on how to use Redux with hooks and redux-thunk

View More Projects @ khems.io