

XINYU ZHAO

☎ 206-245-8834 ◊ ✉ zhaoxy941@gmail.com
200 E Fillmore St, Phoenix, AZ, 85004
🌐 <https://xz941.github.io>

OBJECTIVE

I am looking for 2021 summer internship positions. My research interest lies in machine learning, deep learning, natural language processing and process monitoring. I have been working on several projects related to information extraction with NLP techniques, change detection and anomaly detection.

EDUCATION

B.E in Mechanical Engineering Jilin University, Overall Result: 84.35/100	September 2012 - June 2016
M.S in Mechanical Engineering University of Washington, GPA: 3.72/4.0	September 2016 - June 2018
Ph.D in Industrial Engineering Arizona State University, GPA: 3.71/4.0	August 2018 - August 2022

SKILLS

Programming Languages	Python, Java, MATLAB, R, SQL, C++
Software & Tools	Pytorch, Tensorflow, AllenNLP, CUDA, Git

PROJECTS

Text Mining for aviation accident reports August 2019 - present
Supported by NASA Tempe, AZ

In this project, we build a **Automatic Summarization** model for aviation accident reports to extract key information from the raw reports. After the structured data is collected, we build a **Sequential Event Prediction model** to predict the failure events of the aircraft for proactive management. **Our research has been selected as one of the most outstanding works and we are developing it for an open source python package – PARA-ATM.**

- We solve the sparsity and inconsistency issue from the accident reports using embedding technique. A novel **hierarchical Encoder Decoder** framework is designed to learn the embedding vectors for representing aviation event logs. Our method is able to improve the event log prediction accuracy from 0.76 to 0.82.
- We extract structured information like time, event data from unstructured aviation accident reports. An **attention based Seq2Seq** framework is used for this automatic summarization task. We further design a list of linguistic features according to the domain knowledge for capturing key information in the reports.
- Building a aviation accident simulation module to para-atm – an **open-source software**. The software aims at information fusion for real-time national air transportation system prognostics under uncertainty.

Change and Anomaly detection August 2018 - present
Supported by NSF Tempe, AZ

In this project, we develop suitable **change and anomaly detection** algorithms under certain scenarios. **Our research paper on Adaptive Partially-Observed Sequential Change Point Detection has entered the 2020 INFORMS QSR Best Paper Competition Finalist.**

- Monitoring the heart signal changes. We design physics-based deep learning model for capturing the complex spatio-temporal relationship in heart signals. More specifically, we apply **convolutional-LSTM** and **dilated-CNN** for the **spatio-temporal modeling**.

- Identifying the defect area in materials. **Hidden Markov Model** is used considering the underlying microstructure changes as the transition between hidden states. We treat acoustic emission signals collected from a nanolithography process as the observations of HMM.
- Predicting the aviation trajectory changes with weather data. We apply LSTM and HMM for a time series classification problem.
- Calculating the remaining useful life(RUL) time for machines. We design a **semi-supervised HMM** for the task. The EM algorithm is revised to satisfy the monotonicity constraint according to the nature of RUL prediction. We further apply **Partial Observable Markov Decision Process(POMDP)** for the predictive maintenance task.
- We study the change detection problem under multiple failure mode from a theoretical point of view. We proved two theorem to guarantee the behavior of the proposed change detection algorithm.

EMPLOYMENT

AstrumU

April 2018 - July 2018

Data Scientist Intern

Kirkland, WA

We mainly focus on build efficient algorithms to automatically extract most valuable information for recruiters from unstructured resume data.

- We extract structured information like name, education, skills from unstructured resume. A pipeline is designed for the information extraction task which able to handle resume in different formats. In general, the pipeline solved the problem 1) converting different formats of resume into raw text using **NLTK and tabula-py** 2) extracting key entities like names, education automatically 3) summarizing the skills from the resume automatically using python
- We build up a recommendation pipeline for finding the best candidate for recruiters according to the job requirements. Build visualization for candidate recommendation using **D3.js**

PUBLICATIONS

- Yan, Hao, **Zhao, Xinyu**, Zhiyong Hu, and Dongping Du. "Physics-based deep spatio-temporal metamodeling for cardiac electrical conduction simulation." In 2019 IEEE 15th International Conference on Automation Science and Engineering (CASE), pp. 152-157. IEEE, 2019.
- **Zhao, Xinyu**, Hao Yan, Jing Li, Yutian Pang, and Yongming Liu. "Spatio-temporal anomaly detection, diagnostics, and prediction of the air-traffic trajectory deviation using the convective weather." In 11th Annual Conference of the Prognostics and Health Management Society, PHM 2019. Prognostics and Health Management Society, 2019.
- **Zhao, Xinyu**, Yunyi Kang, Hao Yan, and Feng Ju. "Semi-supervised constrained hidden markov model using multiple sensors for remaining useful life prediction and optimal predictive maintenance." In 11th Annual Conference of the Prognostics and Health Management Society, PHM 2019. Prognostics and Health Management Society, 2019.