# Shuo Yan 闫硕

yanshuok@gmail.com yanshuok.github.io

## Education

- **M.Eng., School of Software, Beijing University of Technology (BJUT), Beijing, China** 09/2016 - 07/2019
  - Overall GPA: 3.5/4.0, top 10%; Advisor: **Bo Liu**
- **B.E., School of Software, Beijing University of Technology (BJUT), Beijing, China** 09/2012 - 07/2016
  - Overall GPA: 3.21/4.0; Major 3.4/4.0

## Awards

- First Prize, Innovation on Technology of BJUT 2019
- Outstanding Master Thesis of BJUT (top 10%) 05/2019
- Outstanding Graduate of BJUT (top 5%) 05/2019
- Second-Class Scholarship of BJUT (consecutive two times, top 20%) 2013~2015
- Third Prize, Innovation on Technology of BJUT (four times) 2017~2019
- Bronze Award, Challenging Cup of BJUT 2015

## Publications

[1] Bo Liu**, Shuo Yan,** Jianqiang Li, Yong Li, Jianlei Lang, Mengchu Zhou, " Study on Prediction of Atmospheric PM2.5 Based on Spatio-Temporal Extreme Learning Machine: Case of Beijing", IEEE Transactions on Big Data (under second review)

[2] Bo Liu**, Shuo Yan,** Jianqiang Li, Guangzhi Qu, Yong Li, Jianlei Lang, Rentao Gu, " A Sequence-to-Sequence Air Quality Predictor Based on the n-Step Recurrent Prediction", IEEE Access 2019 (IF= 4.098, Q1)

[3] Bo Liu, **Shuo Yan**, Jianqiang Li, Guangzhi Qu, Yong Li, Jianlei Lang, Rentao Gu, " An Attention-Based Air Quality Forecasting Method", IEEE, International Conference on Machine Learning and Applications (ICMLA 2018)

[4] Bo Liu, **Shuo Yan**, Huanling You, Yan Dong, Yong Li, Jianlei Lang, Rentao Gu, " Road surface temperature prediction based on gradient extreme learning machine boosting", Elsevier, Computers in Industry 2018 (IF= 4.769, Q1)

[5] Bo Liu**, Shuo Yan,** Huanling You, Yan Dong, Yong Li, Jianlei Lang, Rentao Gu, "Road surface temperature prediction based on gradient extreme learning machine boosting" IEEE, International Conference on Machine Learning and Applications (ICMLA 2017)

[6] Bo Liu, **Shuo Yan**, Jianqiang Li, Yong Li, "Forecasting PM2. 5 concentration using spatio-temporal extreme learning machine" IEEE, International Conference on Machine Learning and Applications (ICMLA 2016)

## Skills

- **Proficient** in Python, Java, Javascript, Html, Spring, SQL, Tensorflow
- **Familiar** with, Matlab, Neo4j, Spark, Linux, Pytorch

## English Proficiency

- **GRE** (April. 27th, 2019): Total: 323 (V: 155, Q: 168) AW: 3
- **TOEFL** (July. 6th, 2019): Total: 102 (R: 29, L: 27, S: 23, W: 23)

## Selected Research & Projects

- **Road Surface Temperature Prediction** 11/2016 - 05/2017

  KEYWORDS: Prediction; Machine Learning; Data Mining

  Published two papers and registered a national patent in China

  Under the guidance of Prof. Dr. Bo Liu

  Used Neural Networks to predict the road surface temperature in the next 24 hours

- **Air Quality Prediction [1]** 07/2016 – 03/2019

  KEYWORDS: Prediction; Machine Learning; Data Mining

  Registered two national-level patents and one software copyright in China; Guided by Prof. Dr. Bo Liu

  Published three papers and do some experiment based on machine learning and deep learning methods

  Design and build a deep learning based air quality prediction system

  Proposed a seq2seq based method to predict air quality

  Significantly reduce the training time of seq2seq and improve the accuracy

- **Improving the accuracy of question searching system [2]** 08/2017 - 09/2017

  KEYWORDS: Cluster; Classification; Text mining; Data Mining

  Identified missing questions of the searching system by machine learning based classifier with labeled data

  Used clustering method to identify the highly frequent missing questions

  The highly frequent missing questions are used to expand the question database

  Built a service which can automatically identify highly frequent missing questions every day

- **Deduplication of question database [5]** 10/2017 - 11/2017

  KEYWORDS: Text similarity, Text Mining

  Use text mining techniques to vectorize the questions

  Use various text similarities to identify duplicated questions

  Build a web service for new imported questions

- **Knowledge point classification [3]** 12/2017 - 05/2018

  KEYWORDS: Text Classification; Text Mining; Data Mining

  Use deep learning and machine learning based methods

  Use reinforcement learning to integrate the word segmentation and classification

## Work Experience

- **Data Mining Intern at Xuebajun, Beijing** 08/2017-08/2018

  KEYWORDS: Software Engineering; Data Mining; Text Mining

  Explored how to enhance the performance of the service about question database by text mining

  Provide some web services based on text mining algorithms

- **Data Mining Intern at China Academy of Science Institute of Automation, Beijing** 07/2019-present

  KEYWORDS: Text Mining, Knowledge Graph, Data analysis

  Build a Knowledge Graph about the company information

  Provide some web services based on text mining algorithms