BRIX LABS

Zhang Jingyao

SUMMARY

Data Scientist with five years of experience in data mining, machine learning and natural language processing(NLP). Achievements include creating ordinal regression models to predict candidate's basic competency score with 83% accuracy. Helped clients to reduce their workload in multiple scenarios by 30% in average. Highly skilled in NLP, machine learning, data visualization, and creative thinking.

SkILL

- Industry Knowledge: Natural Language Processing, Deep Learning, Machine Learning
- ML/DL models: knn, k-means, logistic regression, random forest, xgboost, Seq2Seq, Elmo, Bert…
- Daily use: Python, shell, MySQL, TensorFlow, PyTorch, Scikit-learn, Tableau, Pandas, Numpy, Seaborn,

  Matplotlib, Scipy, AWS services.

WORK EXPERIENCE

AIVALS HRTPS Ltd, AI Lab **Shanghai, China**

Senior NLP Engineer **May 2020 – Present**

Use AI technology to achieve traditional professional HR solutions including:

Developed algorithm for AI assessments which replaced traditional assessment center and was widely used

by Fortune 500 companies in recruitment process.

Developed algorithm for AI coaching and training system which was widely used by top pharmaceutical

companies for medical rep training and development.

Job Recruitment Online Interview Evaluation System (Chinese & English) **Role: project leader**

Project brief

• Closely worked with psychologists and HR consultants in product design, learning the underlying theories

of Behavioral Event Interview (BEI).

• Built the algorithm from scratch based on the theories and best practices in the industry, achieving an

automatic scoring system for 10 basic competencies.

• Built a various of algorithms to support interview evaluation, including detecting STAR structure

(Situation/task/action/result) and irrelevant content.

• Developed algorithm to trigger follow up questions during the interview process, which is unique in

the industry.

Project achievement

• Accuracy reached 70% using ordinal regression model and the relative order information between samples.

• Accuracy increased to 83% and the problem of small set overfitting has been alleviated, by combing tricks like domain data pre-training and sentence-level data augmentation.

• Reduced the problem of inconsistent data labeling using model voting.

• The model's predictions result are highly interpretable.

Online Group Interview Assessment System (Gamified) **Role: Project Leader**

Project brief

• Like most offline group assessment or case study sessions, 4-10 candidates will play against each other in an online game, and the observers (HR) will score candidates' performance according to their language, interpersonal-interaction, facial expressions and voice.

• In specific, algorithm managed to assess candidates' ability of communication and collaboration, their logical thinking, and language ability (Mandarine & English)

• This system reduced workload for HRs and highly increased recruitment efficiency.

Project achievement

• Reduced HRs' working hours by 10% on scoring. This was achieved by building topic thesaurus for each competency using corpus clustering and generating matching template based on syntactic structure.

• Precision reached 77% and recall rate achieved 59% when scoring by semantic matching with all candidates' recall results.

• Produce highlight moment for each candidate to produce a brief and vivid video profile for HRs to review, which has an accuracy of 90%.

DSM AI coaching **Role: Key Development Personnel**

Project brief:

• An assessment tool for evaluating one's coaching ability based on the conversation in one-toone coaching sessions. This tool has been widely used in global pharmaceutical companies to train the career coaches whose job is delivering high quality coaching sessions with medical reps.

Project achievement:

BRIX LABS

• Calculating category probability by semantic similarity matching, and obtain prediction by maximize posterior probability. The double-blind accuracy rate exceeds 90%.

• Predicted scores are used as bonus basis for trainers;

• Build cheating detected methods such as one person playing two roles, reading manuscripts, playing audio

and video, with an accuracy of 95% and a recall rate of 75%;

• Double blind accuracy rate exceeds 90% when calculating category probability by semantic matching and

obtaining prediction by maximize posterior probability.

Zhuiyi Delta Plus INC., AI Lab **Shenzhen, China**

NLP Engineer **Dec. 2017 - May 2020**

Conducted cutting edge research in areas including question answering, information extraction, and language

model, and achieved SOTA results in intelligent Personal Assistant

Chatbot refusal module development and optimization Role: Project Leader

• Identify user intents, reject undefined intents, and classify defined intents normally;

• On the basis of controlling the false rejection rate, fine tuning Bert by multi-task learning, uses the text similarity difference to significantly improve the rejection precision of multi-type negative samples;

International Dialogue System Development Role: Key Development Personnel

• Develop a multilingual versions question and answer system in Arabic, Portuguese, Japanese, Korean.

• Using distillation technology to improve model inference speed.

Midea
Foshan, China

Automation Engineer **Aug. 2014 - Oct. 2015**

Engage in automation-related work and optimize production processes

EDUCATION

University of New South Wales **Sydney, Australia**

Master of Artificial Intelligence and Data Science Mar. 2016 - Dec. 2017

Chongqing University **Chongqing, China**

Bachelor of Electronic Information Engineering Aug. 2010 - May 2014

PATENT

• Multi-user universal question and answer system based on text similarity retrieval. (Patent Number 201910514477.4)

• A training method that enhances the robustness of text classification models. (Patent Number 201910620062.5)