Lei Ding

Email: xiandao.airs@gmail.com Tel: +1 408 219 4062

EDUCATION

University of California, Santa Cruz Aug. 2022 ~ Now
Computer Science Ph.D.: Machine Learning, Large Language Model (**LLM**) & Intelligent Agents Santa Cruz, California, U.S.
Sichuan University Aug. 2004 ~ July. 2007
Master: Applied Computer Technology - Computer Network Sichuan, China
Sichuan University Aug. 2000 ~ July. 2004
Bachelor: Software Engineering Sichuan, China

PUBLICATIONS

- Ding Lei, Li Zhisu, Peng Jian, Shi Xianlin, Wu Wei, Jian Xiaoyu (2007). Deep research on JXME's MIDP protocol and video-sharing framework. Journal of Sichuan University: Natural Science Edition 44(4), 807-811.
- Abena AChiaa Atwereboannah, Wu Weiping, Ding Lei, Sophyanbi B. Yussif, Edwin Tenagyei. Protein-ligand binding affinity prediction using Deep Learning, 2021 18th International Computer Conference on Wavelet Active Media Technology and Information Processing (ICCWAMTIP 2021), 56.
- Wu Weiping, Ding Lei (September, 2022). Digitalization and upgrading of drug design and development by artificial intelligence. Changsha, international peptide drugs and innovation summit, oral presentation.

PATENT& SOFTWARE COPYRIGHT

- Rating of city road segments for taxi hailing based on HANA technology, US Application NO. 13/934,706 | Patent ID 81495268 | Patent Ref 120542US01, China Application NO. 20130269463.3 | Patent ID 82826027 | Patent Ref 120542CN01
- Automatic category assignment and potential topic discovery for products based on Latent Dirichlet Topic algorithm, SAP Patent Invention ID. 83839165
- Simulator of bundle clicking for validating Bandit strategies in A/B testing, US Application NO. 17/547,637 | Patent ID 83839171 | Patent Ref 210412US01
- Reinforcement Learning Model for product recommendation considering balance between product profit and customer interests, US Application NO. 17/556,238 | Patent ID 83848635 | Patent Ref 210416US01
- Software copyright in China for Practice tool v1.3 of GRE analytic writing, registration NO. 2021SR1399061, certificate NO. 8121687

RESEARCH & WORKING EXPERIENCE

University of California, Santa Cruz 2022.08-Present
Computer Science Ph.D. [1]

- Leverage **LLM** to empower intelligent agents for exploring and completing tasks given specific goal, including plan generation and extraction from web search under goal settings (**plan preparation**), building models to predict appropriate actions to be taken (**action prediction model**), and enabling agents to execute actions on target platform (**execution proxy model**).
- **Data collecting and execution simulation**: build up an online platform **MagicWand** to collect training data by allowing annotators to complete tasks on web screen mirror of Android devices, enabling agents to execute plans and collect execution statistics, esp. attributing for predicted actions. build up **apps** to support the aforementioned operations on end-user devices.
- **Action prediction model**: design **multimodal models** that can understand application context via textural and visual information of Android UI and predict the most suitable action given a plan description. Another goal is to transfer this prediction capacity based on UI to boarder platforms, like multiple OS systems and Brower-based systems.
- **Execution proxy model**: design a generic execution layer and implement it on target platforms to allow agents to take execution and give back feedbacks.

[1] Check with blog https://llv22.github.io/orlando.github.io/ about research details and open-source projects

SAP Upscale, SAP Labs 2017.07-2022.08

Senior Data Scientist and Architect

- Conducted convergence analysis of A/B testing of product bundles (products combination) that customers show decayed interests to using Epsilon Greedy, Softmax and UCB1 Bandit Algorithms for customer behaviors' simulation.
- Analyzed product similarity and complement relationship via Latent Dirichlet Allocation Model given the products category tree and the text feature of products.
- Discover and generate product bundle via Collaborative Filtering, Apriori, FP-Growth based on short-term customer interests.
- Generated dynamic product bundles and recommended product items to customers using Deep Reinforcement Learning Algorithms in order to hit the balance between gaining product profit for merchant and fulfillment of customer interests based on product features and short-term customers' behavior data.

SAP Engagement Center on Cloud Infrastructure, SAP Labs 2015.04-2017.06

Senior Software Engineer and Architect

- Conducted service exception discovery model based on Multivariate Gaussian distribution and analyzed efficiency of exception handling in the system.
- Analyzed system bottleneck and optimized the service based on payload statistics and service dependency graph.

Big data application and algorithm optimization in SAP Nanjing Innovation Center, SAP Labs 2011.04-2015.03

Algorithm Lead and Architect

- Optimized the CONOP Program[2] based on Simulated Annealing Algorithm used to determine relative time scale of fossil records, and gave a parallelization proposal based on Monte Carlo sampling: a co-innovation project with Nanjing Institute of Geology and Paleontology[3]
- Designed the Nanjing Smart Traffic Platform, including Origin-Destination analysis, city congestion analysis, dynamic traffic zone extraction, short-term congestion prediction, fake vehicle plate number discovery.
- Analyzed public opinions for Nanjing 12345 hotline by leveraging multi-class classification algorithms: Multilabel K-Nearest Neighbors Algorithm based on label entropy, Euclidean distance and TF-RDF measures, Bayesian classification algorithm with add-on smoothing, Multilabel classification based on Support Vector Machines(SVM).
- Extracted critical public security events based on dependency parsing and analyzed their temporal relationship based on rules.

Platform development and partner toolkit development for SAP Business ByDesign, SAP Labs 2007.05-2011.03

Algorithm and Application Developer

- Design and implement dynamic bird-view for visualization of graphic relationship among business entities.
- Develop Business Object Description Language (BODL) and Advanced Business Scripting Language (ABSL) based on ANTLR, and integrated them with Eclipse and Visual Studio for SAP Partner Development Infrastructure(PDI).
- Integrated Visual Studio plugins and visual editor of User Interface (UI) by an across-AppDomain communication framework that also guarantee the process security via .Net AppDomain isolation.

HONOR & CERTIFICATES

- 2003 Microsoft Innovation Cup - SALT, The school team, 12th place of Mainland China, Touronline(virtual tour online)
- 2010 SAP Excellent Employee，2012 SAP High Potential Employee
- 2013 and 2014 Team Coach of Nanjing Innovation Center in SAP Innovation Competition, 1st place of China Lab
- Tianchi CIMKM AnalytiCup 2018, Cross-lingual Short-text Matching of Question Pairs, 26/1027
- Coursera Andrew Ng's Machine Learning Series(with programming assignment)，refer to certificate
- Coursera Andrew Ng's Deep Learning Series (with programming assignment), refer to certificate
- Coursera Daphne Koller's Probabilistic Graph Model 1，2，3(with programming assignment)，refer to certificate
- Coursera Geoffrey Hinton's Neural Networks for Machine Learning, refer to certificate
- Udacity Michael Littman and Charles Isbell's Reinforcement Learning (with programming assignment[4])
- UC Berkeley CS285 by Sergey Levine，refer to programming assignment

---

[2] CONOP, refer to https://www.paleosoc.org/assets/docs/extended-CONOP-COURSE-NOTES.pdf

[3] Collaboration with Professor Jinxuan, Fan, refer to news in Chinese Academy of Science. About algorithm design refer to SAP China Tech Talk ZE038 delivered by Orlando

[4] It's my private GitHub repository. Once need to access, please contact me with email