个人简历

个人信息

姓名：褚崴

工作经验

- 首席科学家/联合创始人， 2023 年 4 月至今
  无限光年，中国
  职责：研发大模型训练平台，以及在金融医疗等专业领域的应用解决方案。
- 资深总监/研究员，2018 年 7 月-2023 年 4 月
  总监，2017 年 7 月-2018 年 7 月
  认知计算和知识图谱，蚂蚁集团，美国
  职责：带领由 100 余名研究人员和工程师组成的研发团队，研发计算机视觉、自然语言理解和知识图谱技术产品，开发多模态学习解决方案。
- 工程总监，2014 年 11 月-2017 年 7 月
  大规模学习，阿里云，阿里巴巴集团，中国
  职责：研发团队负责人开发分布式机器学习平台，并交付平台产品 PAI 2.0 到阿里云，包括数百个分布式学习的实现集群算法和预测模型的在线服务。
- 主任应用科学家，2014 年 1 月 – 2014 年 11 月
  高级应用研究员，2011 年 5 月-2014 年 1 月
  Microsoft Bing，微软，美国
  职责：Microsoft Bing 的个性化搜索团队负责人，负责研发个性化搜索服务。
- 科学家，2008 年 1 月 – 2011 年 5 月
  雅虎！ 实验室，美国
- 副研究员，2006 年 1 月-2008 年 1 月
  计算学习系统中心，纽约市哥伦比亚大学，美国

教育

- 博士后，2003 年 2 月-2006 年 1 月
  英国伦敦大学学院 (UCL) Gatsby 计算神经科学部（图灵奖得主 Hinton 教授 98 年创立）
  导师：Zoubin Ghahramani 教授（谷歌大脑副院长兼剑桥教授）

- 博士，1999 年 7 月-2003 年 1 月
  新加坡国立大学 (NUS)，新加坡
  导师：Sathiya Keerthi 教授 和 Chong Jin Ong 教授
- 工学硕士，1995 年 9 月-1998 年 1 月
  哈尔滨工业大学/航天部三院，中国
- 工学学士，1991 年 9 月-1995 年 7 月
  哈尔滨工程大学（原哈尔滨船舶学院），中国

荣誉奖励

- 时间考验奖，Seoul Test of Time Award，The Web Conference (WWW)，2023 年
- 时间考验奖，ACM SIGIR，2022 年
- 中国计算机协会 CCF 科技进步二等奖，2022 年
- 中国图象图形学学会技术发明二等奖，2022 年
- 浙江省科学技术进步二等奖，2022 年
- 吴文俊人工智能科技进步一等奖，2020 年
- 最佳演示奖，ACM CIKM，2017 年
- 国家千人，中组部第 12 批长期创新，2016 年
- 浙江省千人计划，2016 年
- 杭州市特聘专家，2016 年
- 最佳论文奖，ACM WSDM，2011 年

学术论文

1. L. Zhang, X. Yan, J. He, R. Li, and W. Chu (2023) **DRGCN:** Dynamic **evolving initial residual for deep graph convolutional networks**, AAAI 2023
2. W. Li, C. Zou, M. Wang, F. Xu, J. Zhao, R. Zheng, Y. Cheng, and W. Chu (2023) **DC-Former: Diverse and compact transformer for person re-identification**, AAAI 2023
3. J. Xu, W. Xu, M. Sun, T. Wang and W. Chu (2022) **Extracting trigger-sharing events via an event matrix**, Findings of the Association for Computational Linguistics: EMNLP 2022
4. Q. Guo, K. Yao, and W. Chu (2022) **Switch-BERT: Learning to model multimodal interactions by switching attention and input**, ECCV 2022: 330-346
5. T.-T. Liang, X. Chu, Y. Liu, Y. Wang, Z. Tang, W. Chu, J. Chen, and H. Ling (2022) **CBNet: A composite backbone network architecture for object detection**, IEEE Trans. Image Process 31: 6893-6906
6. W. Hong, J. Lao, W. Ren, J. Wang, J. Chen, W. Chu (2022) **Training object detectors from scratch: An empirical study in the era of vision transformer**, *in Proc. of CVPR 2022*
7. H. Wang, T.-W. Chang, T. Liu, J. Huang, Z. Chen, C. Yu, R. Li, W. Chu (2022) **ESCM2: Entire space counterfactual multi-task model for post-click conversion rate estimation**, *in Proc. of SIGIR 2022*

8. K. Ji, J. Liu, W. Hong, L. Zhong, J. Wang, J. Chen, W. Chu (2022) **CRET: Cross-modal retrieval transformer for efficient text-video retrieval**, *in Proc. of SIGIR 2022*

9. M. Li, X. Lin, X. Chen, J. Chang, Q. Zhang, F. Wang, T. Wang, Z. Liu, W. Chu, D. Zhao and R. Yan (2022) **Keywords and instances: A hierarchical contrastive learning framework unifying hybrid granularities for text generation**, *in Proc. of ACL 2022*

10. F. Yu, K. Huang, M. Wang, Y. Cheng, W. Chu, and C. Li (2022) **Width & depth pruning for vision transformers**, *in Proc. of AAAI 2022*

11. H. Huang, Y. Wang, Z. Chen, Y. Zhang, Y. Li, Z. Tang, W. Chu, J. Chen, W. Lin, and K.-K. Ma (2022) **CMUA-Watermark: A cross-model universal adversarial watermark for combating deepfakes**, *in Proc. of AAAI 2022*

12. L. Chao, J. He, T. Wang and W. Chu (2021) **PairRE: Knowledge graph embeddings via paired relation vectors**, *ACL 2021: 4360-4369*

13. F. Xu, M. Wang, W. Zhang, Y. Cheng and W. Chu (2021) **Discrimination-aware mechanism for fine-grained representation learning**, *CVPR 2021*

14. W. Hong, P. Guo, W. Zhang, J. Chen and W. Chu (2021) **LPSNet: A lightweight solution for fast panoptic segmentation**, *CVPR 2021*

15. W. Hong, K. Ji, J. Liu, J. Wang, J. Chen and W. Chu (2021) **GilBERT: Generative vision-language pre-training for image-text retrieval**, *SIGIR 2021: 1379-1388*

16. C. Jiang, K. Huang, S. He, X. Yang, W. Zhang, X. Zhang, Y. Cheng, L. Yang, Q. Wang, F. Xu, T. Pan and W. Chu (2021) **Learning segment similarity and alignment in large-scale content based video retrieval**, *ACM MM 2021*

17. K. Chen, W. Xu, X. Cheng, X. Zou, Y. Zhang, L. Song, T. Wang, Y. Qi and W. Chu (2020) **Question directed graph attention network for numerical reasoning over text**, *EMNLP 2020:6759–6768*

18. L. Chao, J. Chen and W. Chu (2020) **Variational connectionist temporal classification**, *ECCV 2020:460-476*

19. X. Chen, W. Xu, K. Chen, T. Wang, S. Jiang, F. Wang, W. Chu and Y. Qi (2020) **SpellGCN: Incorporating phonological and visual similarities into language models for Chinese Spelling Check**, *ACL 2020:871-881*

20. X. Lin, W. Jian, J. He, T. Wang, and W. Chu (2020) **Generating informative conversational response using recurrent knowledge-interaction and knowledge-copy**, *ACL 2020:41-52*

21. F. Xu, W. Zhang, Y. Cheng and W. Chu (2020) **Metric learning with equidistant and equidistributed triplet-based loss for product image search**, *WWW 2020:57-65*

22. S. Wang, B. Zhu, C. Li, M. Wu, J. Zhang, W. Chu, and Y. Qi (2020) **Riemannian proximal policy optimization**, *Computer and Information Science 13(3)*

23. W. Zhang, Y. Cheng, X. Guo, Q. Guo, J. Wang, Q. Wang, C. Jiang, M. Wang, F. Xu and W. Chu (2020) **Automatic car damage assessment system: reading and understanding videos as professional insurance inspectors**, *AAAI 2020:13646-13647 Demonstration Track*

24. W. Huang, X. Cheng, K. Chen, T. Wang, W. Chu (2020) **Towards fast and accurate neural Chinese word segmentation with multi-criteria learning**, *COLING 2020:2062-2072*

25. C. Li, X. Yan, X. Deng, Y. Qi, W. Chu, L. Song, J. Qiao, J. He and J. Xiong (2019) **Latent dirichlet allocation for Internet price war**, *AAAI 2019:639-646*

26. X. Cheng, W. Xu, T. Wang, W. Chu, W. Huang, K. Chen and J. Hu (2019) **Variational semi-supervised aspect-term sentiment analysis via transformer**, *CoNLL 2019:961-969*

27. W. Huang, X. Cheng, T. Wang and W. Chu (2019) **BERT-based multi-head selection for joint entity-relation extraction**, *NLPCC (2) 2019:713-723*

28. W. Sui, Q. Zhang, J. Yang and W. Chu (2018) **A novel integrated framework for learning both text detection and recognition**, *ICPR 2018:2233-2238*

29. T. Yin, X. Deng, Y. Qi, W. Chu, J. Pan, X. Yan and J. Xiong (2018) **Personalized behavior prediction with encoder-to-decoder structure**, *NAS 2018:1-10*

30. J. Yu, M. Qiu, J. Jiang, J. Huang, S. Song, W. Chu and H. Chen (2018) **Modelling domain relationships for transfer learning on retroeval-based question answering systems in E-commerce**, *ACM International Conference on Web Search and Data Mining* (WSDM-11):682-690

31. M. Qiu, P. Zhao, K. Zhang, X. Shi, X. Wang, J. Huang and W. Chu (2017) **A short-term rainfall prediction model using multi-task convolutional neural networks**, *IEEE International Conference on Data Mining* (ICDM)

32. F. Li et al. (2017) **AliMe Assist: an intelligent assistant for creating an innovative E-commerce experience**, *ACM International Conference on Information and Knowledge Management* (CIKM) Winner of the Best Demo Award

33. M. Qiu, F.-L. Li, S. Wang, X. Gao, Y. Chen, W. Zhao, H. Chen, J. Huang and W. Chu(2017) **AliMe Chat: A Sequence to Sequence and Rerank based Chatbot Engine**, *Annual Meeting of the Association for Computational Linguistics* (ACL-55 Short Paper)

34. J. Yang, Y. Chen, S. Wang, L. Li, C. Meng, M. Qiu, W. Chu (2017) **Practical lessons of distributed deep learning**, *Workshop on Principled Approaches to Deep Learning, at ICML*

35. B. Bi, H. Ma, B. Hsu, W. Chu, K. Wang and J. Cho (2015) **Learning to recommend related entities to search users**, *ACM International Conference on Web Search and Data Mining* (WSDM-08):139-148

36. J. Yan, W. Chu, R. W. White (2014) **Cohort modeling for enhanced personalized search**, *ACM SIGIR Conference on Research and Development in Information Retrieval* (SIGIR-37)

37. X. Li, C. Guo, W. Chu, Y. Wang, J. Shavlik (2014) **Deep learning powered in-session contextual ranking using clickthrough data**, *Workshop on Personalization: Methods and Applications, at Neural Information Processing Systems* (NIPS)

38. H. Wang, X. He, M. Chang, Y. Song, R. W. White, W. Chu (2013) **Personalized ranking model adaptation for web search**, *ACM SIGIR Conference on Research and Development in Information Retrieval* (SIGIR-36)

39. R. W. White, W. Chu, A. Hassan, X. He, Y. Song, H. Wang (2013) **Enhancing personalized search by mining and modeling task behavior**, *International World Wide Web Conference* (WWW-22)

40. H. Wang, Y. Song, M. Chang, X. He, R. W. White, W. Chu (2013) **Learning to extract cross-session search tasks**, *International World Wide Web Conference* (WWW-22):1353-1364

41. T. Moon, W. Chu, L. Li, Z. Zheng, Y. Chang (2012) **An online learning framework for refining recency search results with user click feedback**, *Transactions on Information Systems* 30(4)

42. L. Li, W. Chu, J. Langford, T. Moon, and X. Wang (2012) **An unbiased offline evaluation of contextual bandit algorithms with generalized linear models**, *Journal of Machine Learning Research - Workshop and Conference Proceedings 26* (JMLR W&CP-26)

43. P. Bennett, R. W. White, W. Chu, S. Dumais, P. Bailey, F. Borisyuk and X. Cui (2012) **Modeling and measuring the impact of short and long-term behavior on search personalization**, *ACM SIGIR Conference on Research and Development in Information Retrieval* (SIGIR-35) Winner of the Test of Time Award

44. W. Chu, M. Zinkevich, L. Li, A. Thomas, and B. Tseng (2011) **Unbiased online active learning in data streams**, *ACM SIGKDD Conference on Knowledge Discovery and Data Mining* (KDD-17)

45. L. Zhang, J. Yang, W. Chu, and B. Tseng (2011) **A machine-learned proactive moderation system for auction fraud detection**, *ACM Conference on Information Retrieval and Knowledge Management* (CIKM-20 Short Paper)

46. L. Li, W. Chu, J. Langford and X. Wang (2011) **Unbiased offline evaluation of contextual-bandit-based news article recommendation algorithms**, *ACM International Conference on Web Search and Data Mining* (WSDM-04) 297-306 Winner of the Best Paper Award

47. W. Chu, L. Li, L. Reyzin, and R. E. Schapire (2011) **Contextual bandits with linear payoff functions**, *International Conference on Artificial Intelligence and Statistics* (AISTATS-14)

48. T. Moon, L. Li, W. Chu, C. Liao, Z. Zheng and Y. Chang (2010) **Online learning for recency search ranking using real-time user feedback**, *International Conference on Information and Knowledge Management* (CIKM-19 Short Paper) 1501-1504

49. L. Li, W. Chu, J. Langford and R. E. Schapire (2010) **A contextual-bandit approach to personalized news article recommendation**, *International World Wide Web Conference* (WWW-19) 661-670 Seoul Test of Time Award

50. S.-T. Park and W. Chu (2009) **Pairwise preference regression for cold-start recommendation**, *ACM Recommender Systems* (RecSys-03):21-28

51. W. Chu and Z. Ghahramani (2009) **Probabilistic models for incomplete multi-dimensional arrays**, *International Conference on Artificial Intelligence and Statistics* (AISTATS-12):89-96

52. W. Chu and S.-T. Park (2009) **Personalized recommendation on dynamic content using predictive bilinear models**, *International World Wide Web Conference* (WWW-18):692-700

53. W. Chu, et al. (2009) **A case study of behavior-driven conjoint analysis on Yahoo! Front Page Today Module**, *ACM SIGKDD Conference on Knowledge Discovery and Data Mining* (KDD-15 Industry Track):1097-1104

54. R. Silva, W. Chu and Z. Ghahramani (2007) **Hidden common cause relations in relational learning**, *Neural Information Processing Systems* (NIPS-20):1345-1352

55. K. Yu and W. Chu (2007) **Gaussian process models for link analysis and transfer learning**, *Neural Information Processing Systems* (NIPS-20):1657-1664

56. P. K. Shivaswamy, W. Chu and M. Jansche (2007) **A support vector approach to censored targets**, *IEEE International Conference on Data Mining* (ICDM-07):655-660

57. W. Chu and S. S. Keerthi (2007) **Support vector ordinal regression**, *Neural Computation* 19(3):792-815

58. V. Sindhwani, W. Chu and S. S. Keerthi (2007) **Semi-supervised Gaussian process classifiers**, *International Joint Conferences on Artificial Intelligence* (IJCAI-20):1059-1064

59. W. Chu, V. Sindhwani, Z. Ghahramani and S. S. Keerthi (2006) **Relational learning with Gaussian processes**, *Neural Information Processing Systems* (NIPS-19):289-296

60. K. Yu, W. Chu, S. Yu, V. Tresp and Z. Xu (2006) **Stochastic relational models for discriminative link prediction**, *Neural Information Processing Systems* (NIPS-19):1553-1560

61. S. K. Shevade and W. Chu (2006) **Minimum enclosing spheres formulations for support vector ordinal regression**, *IEEE International Conference on Data Mining* (ICDM-06):1054-1058

62. W. Chu, Z. Ghahramani, R. Krause and D. L. Wild (2006) **Identifying protein complexes in high-throughput protein interaction screens using an infinite latent feature model**, *Pacific Symposium on Biocomputing* (PSB-11):231-242

63. W. Chu (2006) **Model selection: an empirical study on two kernel classifiers**, *International Joint Conference on Neural Networks* (IJCNN-06):1673-1679

64. W. Chu, Z. Ghahramani, A. Podtelezhnikov and D. L. Wild (2006) **Bayesian segmental models with multiple sequence alignment profiles for protein secondary structure and contact map prediction**, *IEEE/ACM Transactions on Computational Biology and Bioinformatics* 3(2):98-113

65. W. Chu, S. S. Keerthi, C. J. Ong and Z. Ghahramani (2006) **Bayesian support vector machines for feature ranking and selection**, In I. Guyon, S. Gunn, M. Nikravesh, and L. Zadeh, editors, *Feature Extraction, Foundations and Applications* Springer:403-418

66. W. Chu, Z. Ghahramani, F. Falciani and D. L. Wild (2005) **Biomarker discovery with Gaussian processes in microarray gene expression data**, *Bioinformatics* 2005(21):3385-3393

67. W. Chu and Z. Ghahramani (2005) **Gaussian processes for ordinal regression**, *Journal of Machine Learning Research* 6(Jul):1019-1041

68. W. Chu, C. J. Ong and S. S. Keerthi (2005) **An improved conjugate gradient scheme to the solution of least squares SVM**, *IEEE Transactions on Neural Networks* 16(2):498-501

69. S. S. Keerthi and W. Chu (2005) **A matching pursuit approach to sparse Gaussian process regression**, *Neural Information Processing Systems* (NIPS-18):643-650

70. W. Chu and Z. Ghahramani (2005) **Preference learning with Gaussian processes**, *International Conference on Machine Learning* (ICML-22):137-144

71. W. Chu and S. S. Keerthi (2005) **New approaches to support vector ordinal regression**, *International Conference on Machine Learning* (ICML-22):145-152

72. W. Chu and Z. Ghahramani (2005) **Extensions of Gaussian processes for ranking: semi-supervised and active learning**, *Workshop Learning to Rank* at (NIPS-18):29-34

73. W. Chu, Z. Ghahramani and D. L. Wild (2004) **A graphical model for protein secondary structure prediction**, *International Conference on Machine Learning* (ICML-21):161-168

74. W. Chu, Z. Ghahramani and D. L. Wild (2004) **Protein secondary structure prediction using sigmoid belief networks to parameterize segmental semi-Markov models**, *European Symposium on Artificial Neural Networks* (ESANN-05):81-86

75. W. Chu, S. S. Keerthi and C. J. Ong (2004) **Bayesian support vector regression using a unified loss function**, *IEEE Transactions on Neural Networks* 15(1):29-44

76. W. Chu (2003) **Bayesian approach to support vector machines**, Doctoral Dissertation, *National University of Singapore*

77. K. Duan, S. S. Keerthi, W. Chu, S. K. Shevade and A. N. Poo (2003) **Multi-category classification by soft-max combination of binary classifiers**, *Multiple Classifier Systems* (MCS-04) Lecture Notes in Computer Science 2709 Springer:125-134

78. W. Chu, S. S. Keerthi and C. J. Ong (2003) **Bayesian trigonometric support vector classifier**, *Neural Computation* 15(9):2227-2254

79. W. Chu, S. S. Keerthi and C. J. Ong (2002) **A general formulation for support vector machines**, *International Conference on Neural Information Processing* (ICONIP-09)

80. W. Chu, S. S. Keerthi and C. J. Ong (2002) **A new Bayesian design method for support vector classification**, *International Conference on Neural Information Processing* (ICONIP-09)

81. S. S. Keerthi, et al. (2002) **A machine learning approach for the curation of Biomedical literature - KDD Cup 2002 (Task 1)**, *SIGKDD Explorations Newsletter*, 4(2) Honorable Mention

82. W. Chu, S. S. Keerthi and C. J. Ong (2001) **A unified loss function in Bayesian framework for support vector regression**, *International Conference on Machine Learning* (ICML-18):51-58

美国专利

1. User trustworthiness, US Patent 9519682 B1
2. Determining user preference of items based on user ratings and user features, US Patent 8301624 B2
3. Predicting item-item affinities based on item features by regression, US Patent 8442929 B2
4. Enhanced matching through explore/exploit schemes, US Patent 8244517 B2
5. Character recognition method and device, US Patent 10872274 B2
6. Segmentation-based damage detection, US Patent 10783643 B1
7. Methods and systems relating to ranking functions for multiple domains, US Patent 10019518 B2
8. Personalized recommendations on dynamic content, US Patent 9600581 B2
9. Segmentation-based damage detection, US Patent 11004204 B2
10. Character recognition method and device, US Patent 10872274 B2
11. Online active learning in user-generated content streams, US Patent 99673218 B2
12. Methods and apparatuses for building data identification models, US App. 20180365522 A1
13. Text information clustering method and text information clustering system, US App. 20180365218 A1
14. Multi-sampling model training method and device, US App. 20180365525 A1
15. Question recommendation method and device, US App. 20180330226 A1
16. Feature data processing method and device, US App. 20180341801 A1
17. Text information clustering method and text information clustering system, US App. 20180365218 A1
18. Multi-sampling model training method and device, US App. 20180365525 A1
19. Method and system for training model by using training data, US App. 20180365521 A1
20. Question recommendation method and device, US App. 20180330226 A1
21. Feature data processing method and device, US App. 20180341801 A1