

# Haibin Yu

School of Computing  
National University of Singapore

☎ (+65) 94815539

✉ haibin@u.nus.edu

📄 herokillerever.github.io



## Personal Statement

I am broadly interested in three aspects of machine learning: explanation, optimization and robustness. I believe they are the three pillars for modern machine learning. Firstly, modern machine learning models tend to behave like black-box, to this end, Bayesian machine learning provides a principle way helping us to explain how and why machine learning models behave. Secondly, it is also a commonly known issue for optimizing the machine learning models, interestingly Bayesian optimization also provides a principle way for non-gradient based optimization. Lastly, my work on adversarial machine learning can further enhance the robustness for modern machine learning models.

## Education

- 2014.8–2020.1 **PhD, Computer Science**, *National University of Singapore*, Singapore.  
Supported by Singapore MIT Alliance Research and Technology
- 2017.2–2017.8 **Visiting Student**, *Massachusetts Institute of Technology*, Cambridge.  
Invited and supervised by Prof. Patrick Jaillet
- 2010.9–2014.6 **Bachelor, Mechanical Engineering and Automation**, *Beihang University (BUAA)*, Beijing, China.  
Ranked 6 out of 187 graduates in 2014; 1st Prize of BUAA Scholarship for 3 consecutive years; Outstanding Student of BUAA for 2 consecutive years; Outstanding graduate

## Computer Skills

- Basic Shell, Git, SQL
- Intermediate Linux, Microsoft Windows
- Advanced Python, Matlab,  $\text{\LaTeX}$ 
  - Deep learning libraries: Tensorflow, Pytorch
  - Machine learning libraries: Scikit-learn, GPflow, LightGBM, Xgboost
  - Data processing library: Pandas

---

## Research Interests

- Bayesian machine learning: Gaussian process models, deep Gaussian process models, and Bayesian neural networks.
- Optimization: Bayesian optimization.
- Generative models: generative adversarial networks, variational auto-encoder, and normalizing flows.
- Approximate inference methods: variational inference, Markov chain Monte Carlo.
- Adversarial machine learning; Meta Learning

---

## Publications (\* indicates equal contribution)

- [1] [Haibin Yu](#), Trong Nghia Hoang, Kian Hsiang Low and Patrick Jaillet. **Stochastic Variational Inference for Bayesian Sparse Gaussian Process Regression**. In *Proceedings of the International Joint Conference of Neural Networks (IJCNN-19)*.
- [2] Zhongxiang Dai, [Haibin Yu](#), Kian Hsiang Low and Patrick Jaillet. **Bayesian Optimization Meets Bayesian Optimal Stopping**. In *Proceedings of the 36th International Conference of Machine Learning (ICML-19)*.
- [3] [Haibin Yu](#)\*, Yizhou Chen\*, Kian Hsiang Low and Patrick Jaillet. **Implicit Posterior Variational Inference for Deep Gaussian Processes**. In *Proceedings of the 33rd Conference on Neural Information Processing Systems (NeurIPS-19)*, (200 out of 6743, spotlight).
- [4] Yizhou Chen, Zhongxiang Dai, [Haibin Yu](#) and Kian Hsiang Low. **Recursive Reasoning-Based Training-Time Adversarial Machine Learning**. *Submitted to Proceedings of the 34th Conference on Neural Information Processing Systems (NeurIPS-20)*.
- [5] [Haibin Yu](#), Kian Hsiang Low and Patrick Jaillet. **Semi-supervised Deep Gaussian Process with Deep Invertible Features**. *ongoing*.

---

## Competitions

2018.7 **NUS-MIT Datathon, Medical Imaging**, Singapore, Champion.

Our goal is to develop image analysis tools to enable the automated diagnosis of melanoma from dermoscopic images, particularly, in this datathon, our tasks are:

- Lesion Segmentation.
- Disease Classification.

For the segmentation task, we utilized **DeepLab** to train the model, which finally achieved *mIOU*: 0.84 which is higher than the 2017 challenge. As for the classification task, we implemented two models, namely, **ResNet101** and **Inception Net V4** plus **SVM**. We achieved *AUC*: 0.84 in detecting the melanoma disease and *AUC*: 0.93 in detecting Seborrheic keratosis. For more details, please refer to the datathon website: [NUS-MIT Datathon](#). For our presentation, please refer to: [Presentation](#)



2018.10-2018.11 **Rong360 Financial Algorithm Challenge**, *Multi Financial Scenario Model Training*, Beijing, 44th out of 1298 teams.

The task is to determine whether a user will be overdue or not given different financial products features. Basically this is a binary classification, in the contest, we utilized Random Forest Decision Tree (RFDT) and Xgboost methods as well as multiple feature engineering methods since the feature dimension is high (6000+ with missing values). Finally we achieve an AUC of 0.7264. For more details, please refer to the website: [Rong360](#).

2018.11-2019.1

**JDD-2018 Global digitalization Challenge**, *Population Dynamics*, Beijing, Global Champion (1st out of 1583 teams, 500000 rmb prize).

The task is to use the history changes of mobile device users in several cities, districts and counties transfer between users, rate of mobile communications equipment users and others analog data in different districts and cities, set up reasonable forecast model, make dynamic population change forecast in various districts and counties of the city in the subsequent 15 days. We used ARIMA, Auto-regression and Xgboost methods to solve the problem as well as spending much time on feature extraction. For more details, please refer to the media news: [Chinese](#); [English](#). For materials of the competition, please refer to: [Code and Presentation](#).



---

## Professional Service

Conference **Annual Conference on International Conference on Learning Representations, ICLR**, 2021.

Invited Reviewer

Conference **Annual Conference on Neural Information Processing Systems, NeurIPS**, 2020.

Invited Reviewer

Conference **International Joint Conference on Artificial Intelligence, IJCAI**, 2020.

Invited Reviewer

Journal **IEEE Transactions on Cybernetics**, 2018—Now.

Invited Reviewer

---

## Academics Talks

2019.10 **Why Probabilistic Machine Learning Comes to Rescue**, *Wilmar@NUS Lab*, NUS, Singapore.

2019.11 **Implicit Posterior Variational Inference for DGP**, *AI Seminar*, NUS, Singapore.

2019.12 **Bayesian Machine Learning and Automatic Machine Learning Come to Rescue**, *Sun Yat-sen University Forum for International Young Scholars*, Guangzhou, China.

2020.6 **Bayesian Machine Learning and Automatic Machine Learning Come to Rescue**, *Soochow University Forum for International Young Scholars*, Suzhou, China.

---

## Work Experience

- 2018.4–2018.7 **Advance AI, Data Scientist Intern**, Singapore.  
Work on image re-alignment to improve the accuracy of OCR detection
- Use VGG16 network to predict the rotated degree of images.
  - Top 3 accuracy( $\pm 1$  degree over 360 degrees) achieves 97%
- 2019.4–2019.6 **Huawei Noah's Ark Lab, Research Intern**, Beijing, China.
- Normalizing Flow in Recommendation Systems.
  - Mentor: Dr. Changying Du.

---

## Teaching Experience

- 2016–2017 **Teaching Assistant**, *Artificial Intelligence and Planning*.  
Teaching tutorials, grading assignments, got 4.6/5 feedback

---

## Honors and Awards

### Awards

- 2010 Outstanding Student Cadre of Jilin Province  
2011 Merit Student of BUAA  
2012 3rd.Prize in National Mathematics Contest for College Students  
2012 3rd.Prize in National English Contest for College Students  
2012 Excellent Student Leader of BUAA  
2014 Outstanding Graduate of Beihang University  
2019 Research Achievement Award of NUS  
2019 NeurIPS Travel Award

### Scholarships

- 2011–2013 1st Prize of BUAA Scholarship for 3 consecutive years  
2012 1st.Prize in Timken Scholarship  
2014–2018 Scholarship of Singapore and MIT Alliance Research Technology

---

## Professional Training

- 2018.4–2018.5 **Deep Learning Specialization**, *Coursera*, 100% grade.
- Neural Networks and Deep Learning
  - Improving Deep Neural Networks: Hyperparameter tuning, Regularization, Optimization
  - Structuring Machine Learning Projects
  - Convolutional Neural Networks
  - Sequence Models

---

## Languages

- Chinese **Mother tongue**  
English **Professional**  
Japanese **Basic** *Basic words and phrases only*

---

## Interests

### **Swimming.**

- 1000m: 22 minutes
- 1500m: 35 minutes

### **Running.**

- 1000m: 3 minutes 15 seconds
- 3000m: 12 minutes 20 seconds
- 5000m: 24 minutes 10 seconds
- 10000m: 55 minutes

*Last updated by Haibin on July 30th 2020*