Hung-Yi Lo（駱宏毅）

Assistant Professor
Department of Information Technology and Communication
Shih Chien University, Kaohsiung Campus
200 University Road, Neimen, Kaohsiung, Taiwan, R.O.C.

einslo@gmail.com
 (+886) 918-622-218
http://www.iis.sinica.edu.tw/~hungyi/

RESEARCH INTERESTS

My research interests include music information retrieval, data mining, machine learning, artificial intelligence, multimedia content analysis, pattern recognition, and national health insurance research database.

EDUCATIONS

2007-2013   **National Taiwan University**
Pursuit Ph.D. degree in Computer Science and Information Engineering
- Thesis: Cost-Sensitive Multi-Label Classification with Applications
- Advisors: Shou-De Lin （林守德） Ph.D. and Hsin-Min Wang （王新民） Ph.D.

2002-2004   **National Taiwan University of Science and Technology**
Master of Science in Computer Science and Information Engineering
- Grade average: 93.3   Final rank: 1/46
- Thesis: Incremental Reduced Support Vector Machines
- Advisor: Yuh-Jye Lee （李育杰） Ph.D.

2000-2002   **National Taiwan University of Science and Technology**
Bachelor of Science in Electronic Engineering

1995-2000   **National Kaohsiung University of Applied Science**
Five-year College in Electronics Engineering
- Ranked first eight times and ranked second twice during the ten semesters

SELECTED AWARDS

2013   **IICM Ph.D. Dissertation Award**

2013   **IEEE Tainan Section Best Ph.D. Dissertation Award**

2013   **Technology Transfer to KKBOX**

The automatic music tag prediction method proposed in my IEEE TMM journal paper has become a technology transfer project to KKBOX[1], which is the largest Chinese music service company. The title of the project is "Homogeneous Segmentation and Classifier Ensemble for Audio Tag Annotation and Retrieval". The music tag prediction technique can be applied in music recommendation.

2010   **First Prizes in KDD Cup 2010, of both All Teams and Student Teams**

KDD Cup is an annual worldwide competition on KDD (knowledge discovery and data mining). It is organized by ACM special interest group on KDD, started from 1997. It is now the most prestigious data mining competition. In KDD Cup 2010, the task is to predict student performance on

---

[1] http://www.kkbox.com/

mathematical problems from logs of student interaction with Intelligent Tutoring Systems. Our team is ranked first among 4978 submission entries and wins $7200 U.S. dollars. Competitors include some members of the Netflix Million Prize winner. I am responsible for generating the final submission of our team.

2009	**Third Prize of the Extended Track in KDD Cup 2009**

In KDD Cup 2009, the task is to predict customer behavior (switch provider, buy new products, or buy upgrades or add-ons) from marketing database provided by Orange, a French telecom company. Our team is ranked third among 453 teams and wins $500 Euro dollars. I am responsible for designing new machine learning method.

2009	**First Place in the Audio Tag Classification Task in MIREX 2009**

MIREX refers to Music Information Retrieval Evaluation eXchange. We are ranked first out of twelve submissions from music research institutes and universities.

2008	**Joint Winner of KDD Cup 2008**

The task of KDD Cup 2008 is early breast cancer prediction from X-ray image provided by Siemens Medical Solution, USA. Our team is a joint winner among 128 teams. Competitors include IBM Research Lab. I am responsible for designing new machine learning method and generating final submission.

2008	**National Taiwan University Student Outstanding Performance Scholarship (Academic Category)**

Awards only four students in NTU

WORKING EXPERIENCES

2013-Now	**Department of Information Technology and Communication, Shih Chien University**

Assistant Professor
- Courses: Linear Algebra, Database System, App Programming, PHP Programming and MySQL, Introduction to Computer Science

2013-Now	**ILinkE Company[2]**

Technical Consultant
- Provide technical support related to social network, information retrieval, artificial intelligence, and big data analysis.

2005-2013	**Institute of Information Science, Academia Sinica**

Research Assistant, Speech, Language and Music Processing Laboratory
- Instructor: Hsin-Min Wang Ph.D.

2008-2010	**National Taiwan University**

Teaching Assistant
- Courses: Probability, Data Mining and Machine Learning: Theory and Practice

2003-2004	**National Taiwan University of Science and Technology**

Teaching Assistant
- Courses: Computational Learning Theory and Applications, Electronic Circuit, Engineering Mathematics

2000-2012	**Hall 13, Church in Taipei**

Volunteer of young people counseling and guitar class teacher

---

[2] http://www.ilinke.com/

PUBLICATIONS

- **Journal Papers**

[1] **Hung-Yi Lo**, Shou-De Lin, and Hsin-Min Wang, Generalized k-Labelsets Ensemble for Multi-Label and Cost-Sensitive Classification, *IEEE Transaction on Knowledge and Data Engineering.*(to appear. DOI Bookmark: http://doi.ieeecomputersociety.org/10.1109/TKDE.2013.112)

[2] **Hung-Yi Lo**, Ju-Chiang Wang, Shou-De Lin, and Hsin-Min Wang, Cost-sensitive Multi-label Learning for Audio Tag Annotation and Retrieval, *IEEE Transaction on Multimedia 18(3), pp. 518-529, June 2011.*

[3] **Hung-Yi Lo**, Chun-Min Chang, Tsung-Hsien Chiang, Cho-Yi Hsiao, Anta Huang, Tsung-Ting Kuo, Wei-Chi Lai, Ming-Han Yang, Jung-Jung Yeh, Chun-Chao Yen, and Shou-De Lin, Learning to Improve Area-Under-FROC for Imbalanced Medical Data Classification Using an Ensemble Method, *SIGKDD Explorations, 10(2), pp.43-46, December 2008.* (Invited Paper of KDD Cup 2008 Winner)

- **Conference Papers and Workshop Proceedings**

[4] Tsung-Hsien Chiang, **Hung-Yi Lo**, and Shou-de Lin, A Ranking-based KNN Approach for Multi-Label Classification, *Asia Conference on Machine Learning (ACML 2012)*. (Oral presentation, 27% acceptance rate.)

[5] **Hung-Yi Lo**, Shou-De Lin, and Hsin-Min Wang, Generalized k-Labelset Ensemble for Multi-label Classification, *IEEE Int. Conf. Acoustics, Speech, and Signal Processing (ICASSP2012)*, Kyoto, Japan, March 2012.

[6] **Hung-Yi Lo**, Ju-Chiang Wang, Hsin-Min Wang, and Shou-De Lin, Cost-Sensitive Stacking for Audio Tag Annotation and Retrieval, *IEEE Int. Conf. Acoustics, Speech, and Signal Processing (ICASSP2011)*, Prague, Czech Republic, May 2011.

[7] **Hung-Yi Lo**, Shou-De Lin and Hsin-Min Wang, Audio Tag Annotation and Retrieval Using Tag Count Information, *International Conference on Multimedia Modeling (MMM 2011)*. (Oral presentation, 21% acceptance rate.)

[8] Tsung-Hsien Chiang, **Hung-Yi Lo**, and Shou-de Lin, A KNN-based Ranking Approach for Multilabel Classification, *Workshop on Machine Learning Research in Taiwan: Challenges and Directions (MLRT@TAAI 2010)*. (Oral presentation)

[9] Ju-Chiang Wang, **Hung-Yi Lo**, Shyh-Kang Jeng, and Hsin-Min Wang, Audio Classification Using Semantic Transformation and Classifier Ensemble, *Workshop on Computer Music and Audio Technology (WOCMAT 2010)*. (Oral presentation)

[10] **Hung-Yi Lo** and Hsin-Min Wang, Phoneme Boundary Refinement Using Ranking Methods, *International Symposium on Chinese Spoken Language Processing (ISCSLP 2010)*. (Oral presentation, 29% acceptance rate)

[11] Hsiang-Fu Yu, **Hung-Yi Lo**, Hsun-Ping Hsieh, Jing-Kai Lou, Todd G. McKenzie, Jung-Wei Chou, Po-Han Chung, Chia-Hua Ho, Chun-Fu Chang, Yin-Hsuan Wei, Jui-Yu Weng, En-Syu Yan, Che-Wei Chang, Tsung-Ting Kuo, Yi-Chen Lo, Po Tzu Chang, Chieh Po, Chien-Yuan Wang, Yi-Hung Huang, Chen-Wei Hung, Yu-Xun Ruan, Yu-Shi Lin, Shou-de Lin, Hsuan-Tien Lin, Chih-Jen Lin, Feature Engineering and Classifier Ensemble for KDD Cup 2010, *JMLR Workshop and Conference Proceedings*. (Invited Paper of KDD Cup 2010 Winner)

[12] **Hung-Yi Lo**, Ju-Chiang Wang and Hsin-Min Wang, Homogeneous Segmentation and Classifier Ensemble for Audio Tag Annotation and Retrieval, *IEEE International Conference on Multimedia & Expo (ICME 2010)*. (Oral presentation, 15% acceptance rate)

[13] **Hung-Yi Lo**, Kai-Wei Chang, Shang-Tse Chen, Tsung-Hsien Chiang, Chun-Sung Ferng, Cho-Jui Hsieh, Yi-Kuang Ko, Tsung-Ting Kuo, Hung-Che Lai, Ken-Yi Lin, Chia-Hsuan Wang, Hsiang-Fu Yu, Chih-Jen Lin, Hsuan-Tien Lin, Shou-de Lin, An Ensemble of Three Classifiers for KDD Cup 2009: Expanded Linear Model, Heterogeneous Boosting, and Selective Naive Bayes, *JMLR Workshop and Conference Proceedings, 7: 57-64, 2009.* (Invited Paper of KDD Cup 2009 Winner)

[14] Hsin-Min Wang, Jen-Wei Kuo and **Hung-Yi Lo**, Towards A Phoneme Labeled Mandarin Chinese Speech Corpus, *International Conference on Speech Databases and Assessment (O-COCOSDA 2008)*.

[15] Jen-Wei Kuo, **Hung-Yi Lo** and Hsin-Min Wang, Improved HMM/SVM Methods for Automatic Phoneme Segmentation, *European Conference on Speech Communication and Technology (Interspeech 2007 - Eurospeech)*.

[16] **Hung-Yi Lo** and Hsin-Min Wang, Phonetic Boundary Refinement Using Support Vector Machine, *IEEE International Conference on Acoustics, Speech, Signal processing (ICASSP2007)*.

[17] Yuh-Jye Lee, **Hung-Yi Lo** and Su-Yung Huang, Incremental Reduced Support Vector Machines, *International Conference on Informatics, Cybernetics and Systems (ICICS 2003)*.

- **Tutorial**

[18] Shou-De Lin, **Hung-Yi Lo**, and Cheng-Te Li, Issues of Mining for Heterogeneous Social Networks, *PAKDD 2009*, Bangkok, Thailand, 2009.

RESEARCH EXPERIENCES

- **Audio Tag Annotation and Retrieval**

Audio tags correspond to keywords that people use to portray a music clip in different aspects, such as genre, mood, and instrument. I have proposed a probability ensemble for audio annotation and a ranking ensemble for audio retrieval [12]. The ranking ensemble has won the first place in the audio tag classification task in the MIREX 2009. I have proposed to exploit the tag count information for handling the noisy labeling information in social tagging [7]. I have also developed some novel cost-sensitive multi-label learning algorithms to exploit the tag count and correlation information jointly [1-2, 6].

- **Multi-label Classification Algorithm**

Multi-label classification is an advanced classification problem which aims to predict a set of label simultaneously. I have jointly worked with lab members and proposed a KNN-based learning algorithm for multi-label classification by integrating a ranking model into the neighbor selection [4, 8]. I have also developed a Generalized k-Labelsets Ensemble for both multi-label classification and cost-sensitive multi-label classification [1, 5].

- **Educational Data Mining (KDD Cup 2010)**

At National Taiwan University, we organized a course for KDD Cup 2010. I served as a teaching assistant to help and work with students. We used sparse feature sets trained by logistic regression; and we also considered condensed features and applied random forest for training. Finally I developed a blending model based on regression to combine the results of the students [11].

- **Customer Relationship Management (KDD Cup 2009)**

The challenge of the KDD Cup 2009 dataset is the heterogeneous features, large amount of categorical features and missing values. I have developed a heterogeneous base learner, which is capable of handling different types of features and missing values, and use AdaBoost to improve the base learner [13].

- **Computer Aided Diagnosis (KDD Cup 2008)**

From the KDD Cup 2008 dataset, I observed a patient imbalanced problem that can seriously hurt the generalization ability of a classifier for medical image classification which has not mentioned in the literatures. For solving this problem, I proposed both patient-balanced SVM and patient-balanced AdaBoost based on cost-sensitive learning [3]. Furthermore, I have jointly worked with lab members and proposed a false positive (FP) tolerance method to improve optimal specificity under perfect sensitivity.

- **Phoneme Boundary Refinement**

Phoneme level annotation is especially important for fundamental speech research. I develop phone-transition-dependent classifiers based on support vector machine for refining the hypothesized phone transition boundaries given by the HMM-based Viterbi forced alignment [14-16]. To overcome information losing problem and class-imbalanced problem, I proposed using learning-to-rank algorithms to refine the hypothesized boundaries instead of classification algorithms [10].

- **Early Complications of Radical Prostatectomy (RRP) for Patients with Prostate Cancer**

I have participated in a joint project with Taipei Veterans General Hospital sponsored by the Bureau of Health Promotion. I developed analytical methods for the Taiwan National Health Insurance Research Database. I also develop prediction model of RRP complications based on the sequence data of National Health Insurance record.

CO-WORK WITH M.S. STUDENTS

- **Thesis Topic: Optimizing Specificity under Perfect Sensitivity for Medical Data Classification**
  - M.S. Cho-Yi Hsiao
  - Paper submitted to Pacific Asia Conference on Knowledge Discovery and Datamining
- **Thesis Topic: A KNN-based Ranking Approach for Multi-label Classification**
  - M.S. Tsung-Hsien Chiang
  - Accepted to publish in the Asia Conference on Machine Learning
- **Thesis Topic: Time-dependent Personalized Music Playlist Generation from Implicit User Feedback**
  - M.S. Yin-Hsuan Wei

INVOLVED RESEARCH PROJECTS

- **Next generation Automatic Speech Recognition - Subproject 5: Development of Linguistic and Phonological Knowledgebase for Next-Generation Automatic Speech Recognition**
  - (National Science Council, 2005/8/1-2008/7/31)
  - Project Director: Hsin-Min Wang Ph.D.

- **Next generation Automatic Speech Recognition (Phase II) - Subproject 2: Automatic Labeling and Validation of Speech Databases**
  - (National Science Council, 2008/8/1-2011/7/31)
  - Project Director: Hsin-Min Wang Ph.D.
- **Taiwan e-Learning and Digital Archives Program**
  - (National Science Council, 2007/12/1-2012/12/31)
  - Project Director: Hsin-Min Wang Ph.D.

- 攝護腺癌確診及治療所引起併發症之探討與研究
  - 行政院衛生署國民健康局 2011 年公開招標案
  - Project Director: Wayne Yen-Hwa Chang （張延驊）, M.D., Ph.D.

PROFESSIONAL SERVICES

- **Reviewer**
  - IEEE Transaction on Knowledge and Data Engineering
  - Social Network Analysis and Mining
  - Journal of Information Science and Engineering
  - IEEE Int. Conf. on Multimedia & Expo (ICME)
  - Asia Information Retrieval Symposium (AIRS)
  - IEEE Int. Conf. on Acoustics, Speech, and Signal Processing (ICASSP)
  - International Conference on Neural Information Processing (ICONIP)
  - European Signal Processing Conference (EUSIPCO)

- **International Conference Organizer**
  - Technical Program Committee Member of 2013 Conference on Technologies and Applications of Artificial Intelligence
  - Machine Learning Session Chair of 2013 Conference on Technologies and Applications of Artificial Intelligence
  - Contest Co-chairs of 2014 Pacific-Asia Conference on Knowledge Discovery and Data Mining

- **Invited Talk**
  - 國科會整合型計畫－新世代自動語音辨識技術之研究 2011 暑期講習會

REFERENCES

- Shou-De Lin （林守德）Ph.D.　　　(Ph.D. Advisor)
  sdlin@csie.ntu.edu.tw
- Hsin-Min Wang （王新民）Ph.D.　　(Research Advisor at Academia Sinica)
  whm@iis.sinica.edu.tw
- Yuh-Jye Lee （李育杰）Ph.D.　　　(M.S. Advisor)
  yuh-jye@mail.ntust.edu.tw
- Chih-Jen Lin （林智仁）Ph.D.　　　(Team Leader of KDD Cup 2009 and 2010)
  cjlin@csie.ntu.edu.tw
- Hsuan-Tien Lin （林軒田）Ph.D.　　(Team Leader of KDD Cup 2009 and 2010)
  htlin@csie.ntu.edu.tw