HOANG ANH DAU

Data Mining Lab, WCH 368
Department of Computer Science and Engineering
University of California, Riverside, CA 92521

Email: anh.h.dau@gmail.com/hdau001@ucr.edu
Homepage: http://www.cs.ucr.edu/∼hdau001
Google scholar profile

Summary

I am a final year PhD student specialized in data mining of time series data. I am keen on learning new things and am a fast learner. I am looking for full-time job opportunities in software engineering and data science.

Education

University Of California, Riverside - CA, USA   09/2015–current
    PhD in Computer Science (Advisor: **Dr. Eamonn Keogh**)

RMIT University - Melbourne, Autralia   07/2011–12/2012
    M.S. in Computing

Monash University - Melbourne, Australia   03/2009–06/2010
    M.A. in Communication and Media Studies

Hanoi University - Hanoi, Vietnam   09/2004–06/2008
    B.A. in Foreign Languages

Experience

UCR, Data Mining Lab – *Graduate student researcher*   09/2015–current

- Expand and maintain the UCR Archive of public data sets for time series classification and clustering
- Optimize Dynamic Time Warping's parameter for time series data mining applications
- Design a general framework for time series motif discovery that incorporates user's biases

Mitsubishi Research Laboratories, Data Analytics – *Research intern*   Summer 2018

- Do research in anomaly detection in industrial manufacturing, the robotic insertion task in particular

Mitsubishi Research Laboratories, Data Analytics – *Research intern*   Summer 2017

- Research into anomaly detection in industrial manufacturing, specifically production line of part assembly

RMIT University, School of CS & IT – *Teaching assistant & student researcher*   03/2013–09/2015

- Work as tutor and lab assistant for various courses in Computer Science
- Research into using built-in mobile phone sensors to automatically detect fall and risky driving behaviours

Selected Publications

- **Hoang Anh Dau** and Eamonn Keogh. "Matrix Profile V: A Generic Technique to Incorporate Domain Knowledge into Motif Discovery." *ACM SIGKDD*, Halifax, Canada (2017)
- Germain Forestier et al. "Generating Synthetic Time Series to Augment Sparse Datasets." *IEEE ICDM*, New Orleans, USA (2017)
- **Hoang Anh Dau** et al. "Phone based Fall Detection by Genetic Programming." *ACM International Conference on Mobile and Ubiquitous Multimedia (MUM)*, Melbourne, Australia (2014)
- **Hoang Anh Dau** et al. "Anomaly Detection Using Replicator Neural Networks Trained on Examples of One Class." *Asia-Pacific Conference on Simulated Evolution and Learning (SEAL)*, New Zealand (2014)
- **Hoang Anh Dau** et al. "Optimizing Dynamic Time Warping's Window Width for Time Series Data Mining Applications." *Data Mining and Knowledge Discovery (DAMI)* (2018)
- Chin-Chia Michael Yeh et al. "Time Series Joins, Motifs, Discords and Shapelets: a Unifying View that Exploits the Matrix Profile." *Data Mining and Knowledge Discovery (DAMI)* (2017)

Public Datasets

- **Hoang Anh Dau** et al. "The UCR Time Series Archive." *arXiv preprint arXiv:1810.07758* (2018)
- Anthony Bagnall et al. "The UEA Multivariate Time Series Classification Archive, 2018." *arXiv preprint arXiv:1811.00075* (2018)

Skills

Languages and tools: Python, MATLAB, Shell scripting, Java, Objective C, LaTeX, Git