THE CHINESE UNIVERSITY OF HONG KONG

Department of Computer Science and Engineering

MSc in Computer Science – Project List

Please fill in your **(1) student name**, **(2) student ID** and **(3) SIX** choices which you will consider to register if they are available.

Student name:_____**Student ID:**_____

No.	Supervisor	Projects list	Your Preference					
			1	2	3	4	5	6
1	Prof. CHENG Sheung Chak	Large Scale Machine Learning and Applications						
2	Prof. CHENG Sheung Chak	Advanced Techniques for Building Big Data Applications						
3	Prof. CHENG Sheung Chak	Large-Scale Data Analytics with General-Purpose Distributed Platforms						
4	Prof. CHENG Sheung Chak	A Low-latency system for Distributed Online Learning						
5	Prof. HENG Pheng Ann	Dilated residual networks for automatic 3D biomedical image segmentation						
6	Prof. HENG Pheng Ann	Deep neural netowrk compression for mobile devices						
7	Prof. HENG Pheng Ann	A Virtual Reality based tool for standard plane annotation in volumetric ultrasound						
8	Prof. HENG Pheng Ann	Automatic translation between multiple medical image modalities						
9	Prof. HENG Pheng Ann	Diagnosis of skin cancer using convolutional neural networks						
10	Prof. LEE Pak Ching	Anomaly detection in telecommunication networks						
11	Prof. LEUNG Kwong Sak	Geographical Information System						
12	Prof. LEUNG Kwong Sak	Deep Neural Network for Informatics						
13	Prof. LEUNG Kwong Sak	SEED - Swarm and Evolutionary Enhanced Deep Intelligent System						
14	Prof. LEUNG Kwong Sak	EEG signal analysis for human pressure monitoring by AI techniques						
15	Prof. LEUNG Kwong Sak	Title: Free topic						
16	Prof. SUN Hanqiu	Attention-driven video mashup						
17	Prof. SUN Hanqiu	Video Super-resolution to FHD/4K						
18	Prof. WONG Kin Hong	Smart wearable glasses development						
19	Prof. WONG Kin Hong	Artificial Neural Network research for 3D computer vision						
20	Prof. WONG Kin Hong	Health diagnosis by measuring pulses with a Smart-Watch						
21	Prof. ZHANG Shengyu	Personalized recommendation systems						
22	Prof. ZHANG Shengyu	Reinforcement learning algorithms						
23	Prof. ZHANG Shengyu	Applications of artificial intelligence to finance						
24	Prof. ZHANG Shengyu	Matching and pricing on shared economics						
25	Prof. ZHANG Shengyu	Algorithms for fair allocations						
26	Dr. LAM Tak Kei	Smart kiosk system						

Titles and Abstracts	Supervisor
Large Scale Machine Learning and Applications Description: Machine learning (especially deep learning) has been widely applied in industry today, and distributed machine learning techniques are in high demand in order to process massive data available in this big data era. In this project, you will learn popular and useful machine learning algorithms, and develop a toolkit for large scale machine learning on a general-purpose distributed platform. Such a toolkit will be very useful to numerous applications both in research and in industry. Students who are interested in this project may come and talk to me. My contact info and other info can be found here: http://www.cse.cuhk.edu.hk/~jcheng/ Remark: This project is particularly suitable for students who wish to work as a data engineer/analyst/scientist (all are highly demanded positions in today's market) in the future.	Prof. CHENG Sheung Chak
Advanced Techniques for Building Big Data Applications Description: Big data is one of the hottest topics today, and numerous companies in industry and almost all research institutes/universities are working on big data, especially on topics related to AI and machine learning. In this project, you will learn cutting edge techniques in processing big data, and develop new applications that are valuable to our industrial collaborators. Students who are interested in this project may come and talk to me. My contact info and other info can be found here: http://www.cse.cuhk.edu.hk/~jcheng/ Remark: This project is particularly suitable for students who wish to work as a data engineer/analyst/scientist (all are highly demanded positions in today's market) in the future.	Prof. CHENG Sheung Chak
Large-Scale Data Analytics with General-Purpose Distributed Platforms Description: Data analysis has become a do-or-die requirement for today's business world. General-purpose distributed platforms (e.g., Spark, Hadoop) have been very popular in research and also widely used in industry for large-scale data analytics. In this project, you will learn state-of-the-art general-purpose distributed platforms for large-scale data analytics, and develop new data analytics tools for real applications demanded by our industrial collaborators. Students who are interested in this project may come and talk to me. My contact info and other info can be found here: http://www.cse.cuhk.edu.hk/~jcheng/ Remark: This project is particularly suitable for students who wish to work as a data engineer/analyst/scientist (all are highly demanded positions in today's market) in the future.	Prof. CHENG Sheung Chak
A Low-latency system for Distributed Online Learning Description: Distributed machine learning is popular in both academia and industry. Current solutions, which are mostly based on Hadoop, Spark and specialized machine learning systems, process data in batches. When users implement online learning algorithms with existing systems, they usually suffer from high latency. In this project, you will learn many popular online learning algorithms such as online LSH, ALS, SGD and so on. We will further design and implement a distributed system for low-latency online learning. Students who are interested in this project may come and talk to me. My contact info and other info can be found here: http://www.cse.cuhk.edu.hk/~jcheng/ Remark: This project is particularly suitable for students who aim at pursuing master's or doctoral degrees in the future.	Prof. CHENG Sheung Chak
Dilated residual networks for automatic 3D biomedical image segmentation Automatic segmentation from 3D biomedical images, such as 3D computed tomography (CT) and magnetic resonance (MR) images, is of great significance for clinical practice. The state-of-the-art methods are based on 3D convolutional neural networks	Prof. HENG Pheng Ann

(3D CNNs). However, the 3D CNNs have a large number of parameters and are difficult for training. Nowadays, dilated convolution and skip connections are widely used in computer vision field and achieved success in other recognition problems. In this project, we will study the efficient network architecture for 3D biomedical image segmentation. Primarily, we will investigate the effectiveness of dilated residual 3D CNN for 3D biomedical image segmentation, which combines dilated convolution and skip connection techniques. The specific output of this project includes: 1) Study and compare the performance the different 3D CNNs for 3D biomedical image segmentation. 2) Proposed an efficient segmentation method based on dilated convolution and skip connections. 3) Implement the networks with one of deep learning libraries (e.g., TensorFlow, PyTorch or Caffe) 4) Show the performance of these networks on open challenge benchmark datasets (e.g., whole heart segmentation dataset, brain segmentation dataset).	
Deep neural netowrk compression for mobile devices Deep neural networks (DNN) have achieved success in various computer vision problems, like image classification, object detection, and semantic segmentation. However, these networks are both computationally intensive and memory intensive, and it is difficult to deploy them on mobile devices (e.g., phone) with limited computing resources. Recent efforts toward reducing these overheads involve pruning and compressing the weights of various layers. In this project, we will study and implement an effective model compression method to deploy the CNNs into mobile devices. The specific output of this project includes: 1) Implement an effective model compression technique (e.g., filter pruning) without hurting original network performance. 2) Learn a lightweight deep learning library designed for mobile devices (e.g., TensorFlow Mobile or Caffe2) 3) Develop a specific application (e.g., face detection, object detection) with this technique on mobile devices.	Prof. HENG Pheng Ann
A Virtual Reality based tool for standard plane annotation in volumetric ultrasound Standard plane is important for doctors to monitor fetal growth. Associated with broad field of view, volumetric ultrasound has unique advantages over traditional 2D planar ultrasound in capturing standard planes. However, limited by current tools, annotating standard planes in volumetric data with 2D view is indirect and inconvenient. To leverage the advantage of virtual reality (VR), we propose to develop a VR based annotation tool to greatly facilitate doctors to travel through the real 3D space, so they can annotate the standard planes more quickly and intuitively. With the project, you can learn to handle GUI, VR, volumetric ultrasound data, and the possible future work in machine learning based automatic standard plane detection.	Prof. HENG Pheng Ann
Automatic translation between multiple medical image modalities Different imaging modalities respond to specific tissues. Leveraging the complementary information between different modalities, especially the data representing the similar anatomy, is an interesting topic in medical image analysis field. Modality translation is a possible solution to the problem. In this project, we will focus on the volumetric data of whole heart, including CT and MR. Taking the Generative Adversarial Network (GAN) as a core, we propose to construct the bidirectional channel to translate CT into MR, and also MR into CT. The translation is general and may promote many applications, like volumetric segmentation.	Prof. HENG Pheng Ann
Diagnosis of skin cancer using convolutional neural networks Skin cancer is one of the most common human cancers. As the development of deep learning technique and dermoscopy technique, we can develop image analysis tools to enable the automated diagnosis of melanoma from dermoscopic images. In this project, we shall develop a deep learning method to automatically analyse the skin lesions from the dermoscopic images. We also need to develop a mobile app and integrate the model into it. The related techniques include convolutional neural networks (CNN) and tensorflow.	Prof. HENG Pheng Ann

Anomaly detection in telecommunication networks Anomaly detection is important in many domains. In this project, we will study a class of anomaly detection algorithms specifically for telecommunication networks. We will also implement them on AF-Stream, a distributed stream processing system developed by ourselves. We're also working with Huawei's Noah's Ark Lab to study the real applications of anomaly detection.. Pre-requisites: strong programming skills (in particular C/C++), interest in distributed systems Remarks: Students need to obtain prior approval from me before taking this project.	Prof. LEE Pak Ching
Geographical Information System The project aims at the development of a platform for geographical information on which information collected by users or mobile sensors (carried by volunteers, e.g. in mobile phones or mobile sensors) can be effectively captured and integrated with information amassed from software platforms such as Satellite data, Google earth, Flickers and Twitter for geo-referenced monitoring and decision making. Such platform will prove to be crucial for decision making in the future where ubiquitous sensors will be carried around by people and vehicles in space and time, and information has to be analyzed and courses of actions have to be decided in real time or near real time. Such platform is very instrumental to us for learning about phenomena such as pollution, microclimatic, environmental, and public health conditions of future cities, and to decide on appropriate improvement or prevention measures. This is a real-life project in collaboration with Institute of Future Cities of CUHK. The objective is to be accomplished by achieving the following sub-goals:, Goal 1. To develop a software system for the dynamic and real-time collection of multi-source information and social network from SNS (Social Network Service, e.g. Facebook, Google+) Goal 2. To develop a software platform to assimilate information collected from the mobile devices in Goal 1 with information collected by platforms such as Flicker and Twitter. The platform need to handle multi-sources, multi-temporal and spatial scale data with different visualization for effective decision making and analysis. We will also integrate the GIS with a Data Mining package such as Weka to carry out meaningful analysis and knowledge discovery. In this project, students will need to develop new software and enhance existing platform with new features according to user requirements from the Institute of Future Cities of CUHK.	Prof. LEUNG Kwong Sak
Deep Neural Network for Informatics Deep learning has been popularized in recent years due to the impressive performance in tasks like image recognition, speech recognition, and genome annotation. In some applications, such as human recognition in video and playing Go, and No-Limit Texas Hold'em poker, deep neural network actually outperforms human in terms of speed and capability. In this project, students will be asked to develop a specialized deep neural network system for a data analysis task. You can choose a project either in geoinformatics (e.g. pollution prediction from radar images, geo-annotation from satellite images), bioinformatics (e.g. drug repurposing, disease susceptibility via gene expression analysis) or other Chinese language related applications (e.g. Chinese speech synthesis, Chinese character recognition, Chinese text mining). In this project, you are required to do the following: • To study deep neural network (during summer holiday). • To find and understand the relevant dataset (during summer holiday). • To develop a deep neural network algorithm using either PyTorch or TensorFlow on our department GPGPU computer clusters. • To benchmark your system across a set of problems.	Prof. LEUNG Kwong Sak

• To extend the system to the web interface or mobile interface. This is a research-based project. Therefore, students are expected to be self-motivated. Prerequisite: Strong programming skills in Python and self-motivated to learn new things Equipment: GPGPU computational clusters	
SEED - Swarm and Evolutionary Enhanced Deep Intelligent System There has been overwhelming evidence to show that bees and ants are quite intelligent despite of their tiny brain. Ants can farm and cultivate crops while bees can understand abstract concepts, such as direction and combination. Furthermore, based on the Darwin's theory of evolution, the emergence of intelligence may be evolved from more simplistic ancestors naturally over time in order to overcome the challenges in different living environments. In this project, we want to create a swarm of intelligent beings via evolutionary computation. We called the whole system as the Swarm and Evolutionary Enhanced Deep (SEED) intelligent system. In this intelligent system, the swarm of intelligence is represented using a population of deep neural networks. Deep neural network is by far one of the most generic intelligent programs. These deep neural networks will collectively evolve to adapt to the fitness functions (i.e. the living environments). Some examples of evolution include optimizing the parameters, choosing the right optimizers or selecting suitable network structure responding. Since evolution takes time, a part of the system will be accelerated on multiple FPGAs. We hope that by applying evolutionary approach in a collaborative manner, we can cultivate even more intelligent programs in the SEEDI system. Lastly, the system will be apply to solve some real world problems (e.g. image recognition). Students will be asked to implement the SEED intelligent system. You are required to do the following: • To study deep neural network (during summer holiday). • To understand the library (see the list below) depending on the given FPGA hardware (during summer holiday). • To deploy a deep neural network accelerated by FPGAs. • To set up a (computer vision) benchmark for comparison across different systems. • To evolve the parameters of the deep neural network. • To evolve the structures of the deep neural network. • To evolve the training behaviors of the deep neural network. (optional) This is a research-based project. Therefore, students are expected to be self-motivated. Prerequisite: Strong programming skills in C++ and Python, and self-motivated to learn new things. Experience in OpenCL development is a plus Equipment: GPGPU computational clusters, FPGA boards	Prof. LEUNG Kwong Sak
EEG signal analysis for human pressure monitoring by AI techniques Electroencephalography (EEG) is used to record electrical activity of the brain. Many diseases can be diagnosis by EEG, like sleep disorders, coma and brain death. In this project, students are required to study how to use EEG signal to indicate pressure (stress). A wireless EEG sensor, EMOTIV Epoc+ 14 channel EEG sensor, will be used in this project, so that people can use the sensor at home to check their mental status. The goal of this project is to use machine learning to analyze how EEG signal related to pressure. We will be collaborating with experts form Psychology Department of CUHK. Equipment: EEG measuring device will be provided	Prof. LEUNG Kwong Sak
Free topic	Prof. LEUNG Kwong Sak

Any project proposed by student or agreed between the student and myself.	
Attention-driven video mashup With the popularization of mobile devices, people are taking videos of the same events everywhere. Even these videos may be uploaded and crowd sourced, the browsing experiences are very limited due to monotonous viewing and visual redundancy. The aim of this project is to develop the system that can automatically combine such videos and generate one video mashup, which summarizes the important events recorded by the videos. The system is able to determine which event is important by analyzing the joint attentions, select the efficient views according to the video quality, and automatically zoom in if necessary during the video browsing. Pre-requisite: C/C++ programming, OpenCV, Computer Vision	Prof. SUN Hanqiu
Video Super-resolution to FHD/4K Full High Definition (FHD) display is the majority nowadays, and 4K will be popular for rich-media devices. To display old video content in Standard Definition (SD) or HD resolution on these displays, interpolation used to up-scale the content will cause poor visual quality such as blurriness. In this project, we will investigate the feasibility of using image super-resolution technique to scale up the video contents. We will extend the sparse representation based image super resolution method to video super resolution problem. We will further utilize redundancy of adjacent images and update dictionary to improve sparse representation based reconstruction. Based on our previous work, we will test the extended approach on real-scene video, and measure the performance for video super resolution tasks. Prerequisite: image/video processing, C/C++ programming	Prof. SUN Hanqiu
Smart wearable glasses development Wearable glasses such as the google glass draw a lot of attention from the public because it provides novel applications to users. For example, it can display texts and information directly to users onto the glasses they are wearing. At CUHK, we designed a novel wearable glass hardware that enlarges the display area and can be built at low cost. The function is more versatile such that one can see through the display so that texts and graphics can be overlaid onto the real scene. This project is to develop software applications such as text recognition, language translation or face recognition applications that can be incorporated into our wearable glass system. The project is suitable for those who are interested in system development and the creation of new products and ideas. We expect the system has high commercial potential and can be turned into a successful commercial product in future. The hardware design of our system can be found at http://www.cse.cuhk.edu.hk/~khwong/proj/msc_khw_1718.pptx	Prof. WONG Kin Hong
Artificial Neural Network research for 3D computer vision Normally Artificial Neural Networks such as the Convolution or Deep Neural Networks can perform well in object recognition tasks, etc. We suggest we can use the networks for 3D reconstruction of models or camera pose tracking as well. The results can be useful in virtual reality systems and other applications such as automatic driving. Traditionally, a network takes the raw images as the input, our new idea is to feed the 2D feature tracking data to the network rather than the raw images to enable better results. We hope this can also enhance the efficiency of the network in terms of training and recognition. Our approach is not just preparing data for large scale training but rather to investigate ways to redesign a new network architecture for better performances. References are: http://www.cse.cuhk.edu.hk/~khwong/proj/msc_khw_1718.pptx, http://deeplearning.net/tutorial/lenet.html, https://www.tensorflow.org/	Prof. WONG Kin Hong
Health diagnosis by measuring pulses with a Smart-Watch It is believed pulse patterns reflect the health status of a person. Traditional, doctors of Chinese medicine diagnose diseases by	Prof. WONG Kin Hong

measuring patterns of patient's pulses to find out the problems and cure. It is found that latest Smart-Watches can provide accurate measurements of such pulse data. By analyzing pulses using machine learning and other artificial intelligence techniques, it is hoped that it will provide enough information for accurate disease classification. The student will work on the development of machine algorithms tailored for this application.	
Personalized recommendation systems. Recommendation systems have a wide range of applications in e-commerce, advertising, search engine, financial services, etc, Many emerging scenarios call for new recommendation algorithms. In this project, you'll first learn some basic recommendation methods, then work with me an a PhD student to try them on real-world data sets. You'll also get chances to design algorithms and fine-tuning them on real AppStore, news recommendation, and job recommendation systems. There is a large demand in recommendation system algorithm designer, and the knowledge and techniques learned in this project will make you well equipped for related job opportunities.	Prof. ZHANG Shengyu
Reinforcement learning algorithms In many practical scenarios, the data items comes in a stream and we need to make sequential decisions under uncertainty. With no or very limited knowledge about the future data, we need to explore around and at the same time aim to maximize the overall performance. Multi-Armed Bandit (MAB) and Markov Decision Process (MDP), two fundamental models in reinforcement learning, address this general problem. In this project, you'll read some of the simple but elegant algorithms with surprisingly good performance, as well as some recent development of the area, and then work with me an a PhD student to conduct some simulations to test their performances. If you have time and good ideas, you are also welcome to try to design new algorithms or heuristics to improve existing ones, and apply the algorithms on real-world datasets.	Prof. ZHANG Shengyu
Applications of artificial intelligence to finance Artificial intelligence (AI) has been successfully applied to many IT related companies, but much less to financial industry. What are typical applications of AI in finance and what will be the major next 5 to 10 years? In this project, you'll first survey the existing AI techniques applied in finance, and then focus on one or two of them to work with me an a PhD student to make further investigations. You'll have a chance to test your developed algorithms or methods on some real financial data.	Prof. ZHANG Shengyu
Matching and pricing on shared economics Consider the scenarios of shared taxi. At any moment, the taxi company receives a large collection of requests and taxi locations, and needs to assign (or recommend assignment) the taxis to requests. The requests may be paired up, meaning that two passengers share the same taxi. How to assign taxis and requests is an optimization problem, and how to price the passengers and charge the drivers is an interesting economical problem. In this project, you'll work with me an a PhD student to investigate this topic and try to design algorithms for different contexts. Your algorithms will then be tested on real datasets to evaluate its performance.	Prof. ZHANG Shengyu
Algorithms for fair allocations. Scenario 1: You are in a party with friends and it's time to cut the cake. You like the chocolate flag and some one else hopes to get a strawberry. In general, different people may have different preferences on different parts of the cake. Is there a way to divide the cake so that each person feels that s/he gets the best part? Scenario 2: You just arrived at HK and need to find an apartment with two other friends. After finding one, you guys face the problem of who takes which room and pays how much. Is there a way to assign the room and set their prices, so that each person feels that s/he gets the best deal in the assignment? Surprisingly, the answers to both questions are Yes! These are just two sample problems in a fascinating area called fair division or fair allocation. Do you like to know more stories, see how the fair allocation is actually done by an intriguing procedure, and even try to design some algorithms of your own? In this project, you'll read some classic problems in the area, to work with me an a PhD student to	Prof. ZHANG Shengyu

implement certain algorithms, to look for new applications, and if you have time and good ideas, also to try discovering new algorithms.	
<u>Smart kiosk system</u> This project aims to create a smart kiosk system. There are many ways to make the system smart: e.g. users may use speech (Cantonese) commands to do certain tasks; the system can analyse and figure out the usage patterns automatically via some machine learning techniques. Students can choose to make (certain parts of) the hardware and/or (certain parts of) the software.	Dr. LAM Tak Kei