Personal Statement

- Applicant: **Kaikai Zhao**
- Proposed Department and Degree: **Ph.D. in Computer Science**
- Semester and year that the student is applying for: **Fall Semester, 2018**

Introduction and Education Background

As an applicant to the doctoral program in Computer Science. I am very glad to have an opportunity of introducing my research interests and experiences to you.

At present, I am a joint civilian postgraduate student in Department of Computer Science, National University of Defense Technology & Naval Aeronautical University. I have taken some machine learning related courses during the past year in NUDT. Prior to NUDT, throughout four years' undergraduate life, I have completed some fundamental courses related to mathematics, physics, programming and information theory.

Research Interests and Specific Goals

My research interests are as follows:

- **Large-scale machine learning**
- **Pattern recognition**
- **Computer Vision**
- **Biometrics**
- **NLP**
- **Deep learning**

I have conducted research on machine learning and submitted our work *Large-scale k-means clustering via variance reduction* to Neurocomputing two months ago. In this paper, we proposed a new method to accelerate k-means by using variance reduction technique denoted as VRKM. Moreover, VRKM++, the variant of VRKM, does not have to compute the batch gradient, and is more efficient. Besides that, I have written a paper *Multiple kernel k-means clustering with late fusion* which is being revised by [Xinwang Liu](), my research mentor. I am the main contributor for this paper, including theoretical derivation(three-step alternate optimization) and empirical evaluations(matlab implementation). In this paper, an effective method is proposed for multiple kernel k-means clustering which integrates the clustering results from multiple views into an optimal one.

Recently, I mainly focus on large-scale machine learning. In terms of large-scale tasks, since kernel matrix imposes quadratic dependence on the number of samples n, traditional kernel method incurs large computational and storage costs, which prevents its applications in large-scale settings. Therefore, research on large-scale kernel learning is meaningful and I read some papers related to large-scale kernel learning in NIPS and ICML Conference proceedings in the last few years.

After the literature survey, I found that Nyström method and random Fourier features are most popular, which were proposed to overcome the computational challange and scale the kernel method to large-scale datasets. Also, I reimplemented the ideas in the papers that I am very interested in. During the past month, to explore the differences between kernel approximation methods and neural

networks, I am implementing large-scale kernel learning via Keras by the fact that the approximated kernel features can be interpreted as the output of the hidden layer in a shallow neural network. Specifically, the bottom layers of these networks are random basis functions which are not adapted during training. At present, I only focus on Gaussian kernel and Arc-cosine kernel. Specifically, the approximated kernel features above are based on random Fourier features proposed by Ali Rahimi and Benjamin Recht in their NIPS 2007 paper.

These days, deep learning is very popular and it has achieved great achievements in many applications, like computer vision, natural language processing, medical image analysis, etc. Also, I am very interested in deep learning and master two deep learning frameworks, i.e. Keras and Mxnet. It is interesting to note that our group members have implemented Hinton's **CapsuleNet** with Keras and the main contributor of implementation is Xifeng Guo, one of our group members. If you are interested in the work mentioned above, they are all available on my Github account.

Meanwhile, I have developed a long-term professional objective: the advancement of large-scale machine learning technology. In my view, we should apply our advanced machine learning algorithms to practical applications, which is also our ultimate goal as scientific researchers.

My thinking about the capsule network

As we all know, CapsuleNet has been very popular before its corresponding paper was published on NIPS 2017. Deep learning researchers know the fact that **pooling operations** make deep neural networks lost space relationships of objects in images. Obviously, there are no **pooling layers** in CapsuleNet, which makes it has the capacity to capture the **space relationship**. In fact, there are many relationships which could(or should) be captured, such as angle relationships. Unfortunately, AI researchers have not yet developed a powerful deep network to capture those meaningful latent relationships so far. Why? In my view, designing a neural network is not difficult, but how to convert that net into a **tractable** problem is not an easy task. Even so, I believe it is promising in this area.

Why do I want to do research and get a PhD degree?

I guess this question is what you care about most. Yes, the answer to this question is very important for me as well.

First, I have to say that my work experience helped me realized I wanted to know more. I want to know why some algorithms produced good results and others do not. Also, I want to propose my *own* algorithms. To be honest, if I am engaging in an interesting work, such as work related to large-scale machine learning, I would stay in the lab all day until I got a satisfying result.

Second, I love machine learning. Machine learning is an intersection of various areas, including mathematics, computer science, and neuroscience and cognitive science. It's a wonderful thing to solve a tricky problem with a simple ML algorithm. For large-scale problems, traditional kernel methods are typically expensive in both memory and time, but we can overcome this difficulty by approximating kernel functions through random features. All the approximation methods are good, because to some extent, they reduce the space and time complexity without a significant effect on the precision.

Third, in my view, a PhD degree is indispensable for scientific researchers and it will help me improve my academic skills and delve deeper in areas of interest. The process that I take a few years to get a PhD degree is very important for my future research career. I wish to be an **independent** researcher with **critical thinking** after I finish my PhD program.

Last but not least, if you feel that I am more suited for some other areas, any of your suggestion would be highly appreciated. I also believe that I possess the capacity for hard work and the motivation required to perform well in a PhD program. Since my parents do not have jobs in rural areas and my economic resources permit me to partially fund myself for the duration of graduate course, I request you to take a tuition free offer into account. As for accommodation and living costs, it would be appreciated that you could fund my living expenses, otherwise any form of financial aid, such as teaching and research assistantships, are highly acceptable, which can also open my mind and improve my social skills.

Life and Learning Experiences

I am the first in my family to attend college and I have a younger brother in my family. When I was young, my family lived in rural areas and my parents did odd jobs to make money. From I was seven years old, I began to cook meals for my family in order to reduce my parents burdens and save their time. Then they had more time to earn more money to pay our tuition. Today, it is incredible to look back upon that hard experience which made me grow up a lot. Now I am an adult and I am grateful for all people who have helped us. Hence, Long Gao, my group member in the lab, and I decided to launch a co-reading program to help distant children read books, to answer questions and doubts, grow together and change their fate through reading. Here is our website: http://www.1book1dream.com.

In the third year of my undergraduate life, I worked on a NSFC project with Dr. Yang, chief researcher of Information Processing Research Laboratory. During that period, I developed research skills and mastered programming skills with **MATLAB, C, C++ and Python**. I love programming, because I can fulfill my ideas by several simple commands. That's really amazing.

In addition to academic work, I love team sports, such as football. For me, playing football is a good stress-reliever. In the last year of my undergraduate life, as the captain of our College team, I led our team to the second place in Changchun University Football League. That's a memorable experience.

After I received my Bachelor Degree, I worked as an engineer in Longkou Longpeng Precision Copper Tube Co. Ltd, a branch of Golden Dragon Precise Copper Tube Group Inc., from Aug. 2013 to Jan. 2016. My main tasks related to that job position were: writing, improving and maintaining technical program on lathe control system.

With all sincerity, I hope you can take my application into serious consideration, and I shall appreciate you if you grant me an opportunity for future studies.

Wishing you and yours a happy 2018.

Sincerely yours,

Kaikai Zhao