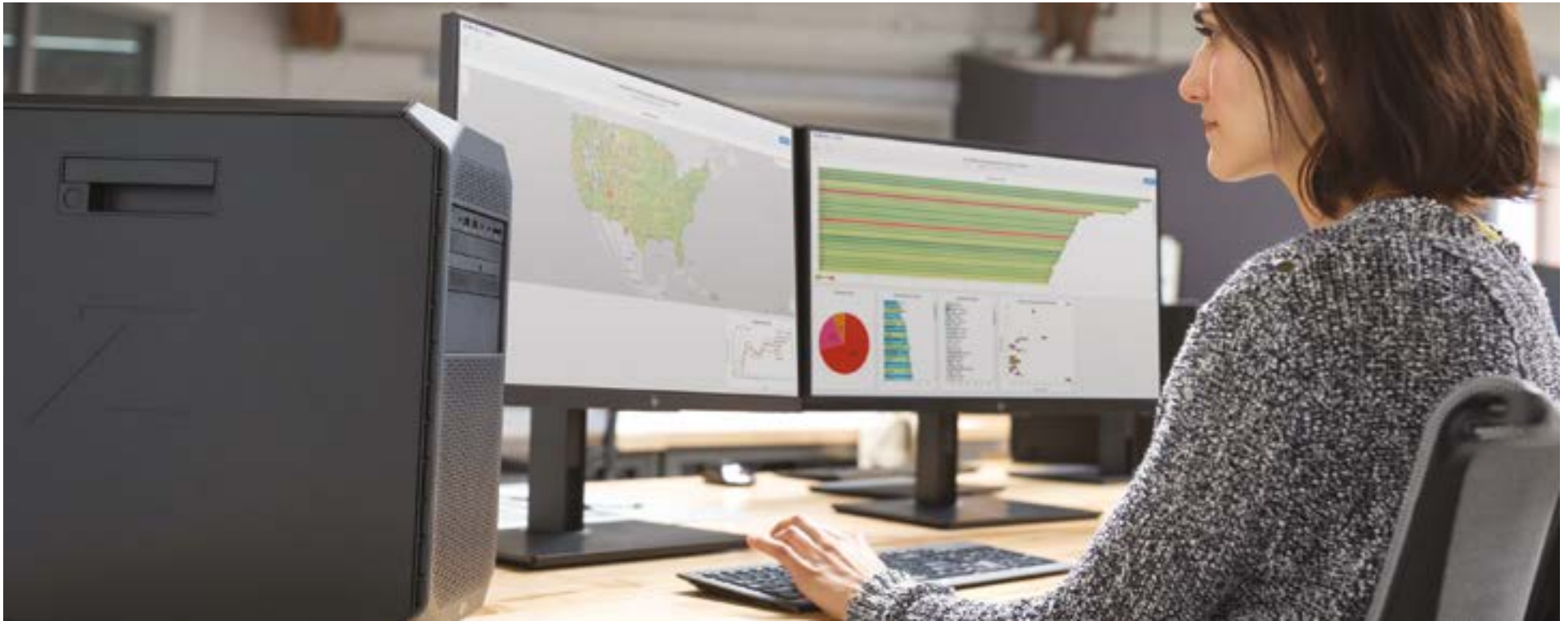




YUANHAO WU: DRIVEN, FOCUSED, CONSTANTLY LEARNING

Z by HP inspires Kaggle Grandmaster to contribute to a medical research program and explore the potential of AI





Few can say that they conclusively decided what they want to do later in life as children. But Yuanhao Wu knew as early as primary school that he wanted a career that involved working with computers. It was there that he learned to program in Logo, a fun children's programming language for drawing on a computer, and that first taste of technology eventually led him to a career in data science. At Tsinghua University, he majored in mechanical engineering, minored in computer science, and obtained a masters in mechatronics. After building

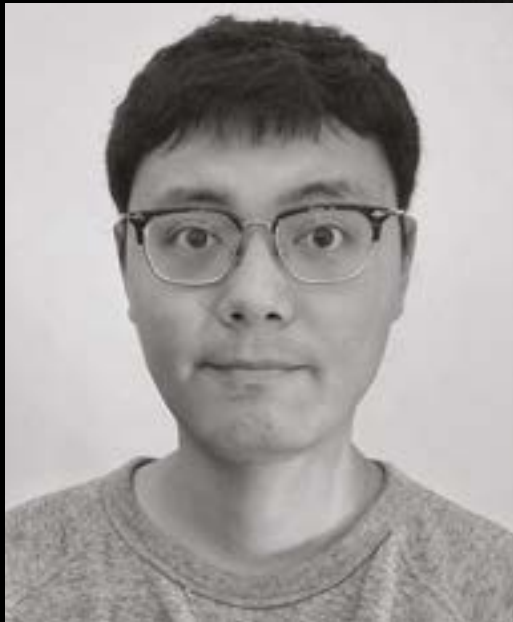
robots in college, Wu realized he was more interested in software than hardware and found he had a special interest in artificial intelligence and AI-powered analytics.

Data science and artificial intelligence really help people, Wu explains. "For example, in China there are many workers doing tedious jobs for manufacturers, such as checking for product defects on an assembly line. The work is really boring and affects their health. Artificial intelligence can help perform that same task and reduce their workloads."

AT A GLANCE

YUANHAO WU, SHANGHAI, CHINA

- Algorithm Engineer with News Break
- Bachelor of Science, (Major) Mechanical Engineering Tsinghua University
- Bachelor's (Minor), Applied Computer Science Tsinghua University
- Master of Science (M.S.) Mechatronics Tsinghua University
- Kaggle Competitions Grandmaster
- Participated in 24 competitions in four years
- Enjoys badminton, reading, movies, and photography



As a Z by HP Global Data Science Ambassador, Yuanhao Wu's content is sponsored and he was provided with HP products.

He looks forward to what the future holds for artificial intelligence. "Right now, the algorithms are only able to perceive what we teach them. Machines can 'see,' but they don't

Kaggle competitions—and that he learns very quickly—has increased his self-confidence. It has also helped him hone his skills and made him more patient. "Kaggle has a very good

"THE HP ZBOOK STUDIO MOBILE WORKSTATION HAS SUCH A POWERFUL CONFIGURATION—THE CPU, THE GPU, AND THE LARGE RAM—AND I HAVE NEVER WORKED WITH SUCH POWERFUL COMPUTERS AS THE HP Z4."

Yuanhao Wu, Z by HP & NVIDIA® Data Science Global Ambassador

understand what humans are talking about, or nuances of humor, or what knowledge really is. The next step is to teach machines to understand, to give them cognition so they can cooperate with humans."

A community of sharing and collaboration

When Wu isn't busy at his day job as an engineer with News Break, a U.S. news aggregator dispensing information about local happenings, he spends time on Kaggle, which he discovered on the Internet after college. "When I first explored it, I thought to myself, this is a wonderful community. I can do a lot of things here, and there are so many helpful people waiting to share their insights and skills. It helped me determine I wanted to stay in this arena."

He spends one or two hours on Kaggle during the week, and when he engages in competitions on weekends, he spends roughly four to six hours a day on the platform. Knowing what he can achieve in the difficult

return," he says. "I invest the time and I get something back. There are other platforms like Kaggle in China that host competitions, but people are not as active in those forums. They don't share their insights and technology like Kagglers."

One misconception about Kaggle he's noticed is that it's all about competitions. "There's much more," he says. Members share technology and what they've learned, and there are also courses for continuing education. He thinks industry should care about Kaggle because many of its members are focused on artificial intelligence and enterprises can learn from them. Also, the forum houses a great repository of ideas and models that could potentially be of use to them.

Wu suggests that both junior and senior data scientists can benefit from being on Kaggle, but he says it's especially useful for junior data scientists because they can learn basic knowledge and tips. He would advise new members not to care too much about



SOLUTION HIGHLIGHTS:

HP Z4 G4 DESKTOP WORKSTATION

- Intel® Xeon® 2295 3 GHz 18C - CPU
- NVIDIA® Quadro® RTX™ 6000 - GPU
- 128 GB (4 x 32 GB) 2933
- Z by HP Data Science Software Stack

Z38C DISPLAY

- 37.5-inch diagonal curve
- 3840 X 1600 at 60 Hz resolution

HP ZBOOK STUDIO G7

- Approx: 4 lbs/2 Kilos
- i9 – 10885H processor
- 32GB RAM
- NVIDIA® RTX 5000 (16 GB VRAM)
- 15.6-inch display
- Ubuntu 20.04
- Z by HP Data Science Software Stack

PRE-LOADED, POPULAR DATA SCIENCE TOOLS, TURN-KEY & READY OUT OF THE BOX:

Z by HP Data Science Software Stack

Learn more:

Z by HP Data Science Brochure

their competition rankings, and instead focus on what they learn and how they can apply it to their fields. “Don’t compete for results and the results will come to you,” he advises.

Power, ease of use, and storage

Data science is used everywhere; for example, in building self-driving cars that according to some experts will reduce the number of car accidents, to deciding airline flight schedules so there are fewer delays. It’s also important in logistics, such as in UPS’s delivery routes and times, and it is used to find predictive patterns, such as in people’s buying or movie-watching habits. In examples like these, data scientists use tools like modeling and experiment with different methods, like Wu does on Kaggle.

As a Z by HP & NVIDIA® Data Science Global Ambassador, Wu has been using the HP ZBook Studio Mobile Workstation and HP Z4 Desktop Workstation while on Kaggle. He is familiar with HP laptops from using them in college. At SAIC, his previous employer, he used an HP Z4 Desktop Workstation.

“The HP Z Book Studio Mobile Workstation has such a powerful configuration—the CPU, the GPU, and the large RAM—and I have never worked with such powerful computers as the HP Z4,” Wu explains.

He especially appreciates the preloaded Z by HP Data Service Software Stack in the HP Z solution. “The preloaded software stack on Z by HP allowed me to start working right after I turned on the computer the first time,” he says. He saw many commonly used libraries and other software, so there was no extra effort needed, adding, “That was really convenient.”



“I LOVE THAT THE Z TEAM IS ALWAYS LOOKING FOR FEEDBACK FROM DATA SCIENCE PRACTITIONERS, TO KEEP THEIR FINGERS ON THE PULSE OF WHAT’S HAPPENING IN THE FIELD. I HAVE NEVER BEFORE FELT THAT A COMPANY BUILT A PRODUCT FOR ME. THE Z BY HP WAS BUILT FOR ME.”

Yuanhao Wu, Z by HP & NVIDIA® Data Science Global Ambassador



“Enterprises will appreciate the HP Z solution because it’s such good quality. Developers and others can focus on their roles and let the HP solution do its part,” he says. Also, he adds, HP is well known for its service.

Wu finds the compact size of the HP Z solution appealing as well. Houses in China are expensive, so many people live in smaller

homes and don’t have space for bulky equipment, he explains.

Z by HP in action

For Kaggle competitions, Wu depends on the power of the HP ZBook Studio Mobile Workstation and he is also able to run technical experiments with it. “I try some newly released models or software, and because the

configuration is so powerful, I can try almost anything I want to,” he says.

One competition where it helped tremendously was the “Hacking the Kidney” computer vision code competition, which is helping the Human BioMolecular Atlas Program (HuBMap) develop “a framework for mapping the human body at a level of glomeruli functional tissue units.” It begins by mapping

the human kidney at single cell resolution. Wu and others developed segmentation algorithms to identify certain elements in photos of kidneys. Competitions like this which are related to the HUB program can result in insights that help pharmaceutical companies developing therapies, and other researchers.

Using the Z4 workstation in this particular competition, he was able to slash the time spent



KAGGLE: LEARN, COMPETE, CHANGE THE WORLD

- Over 6 million registered Kagglers solve data science problems, gain access to powerful tools and resources, and compete for prizes.
- There are five tiers: Novice, Contributor, Expert, Master, and Grandmaster. Currently, there are 211 Grandmasters.
- Kaggle has run hundreds of competitions, from improving gesture recognition to improving the search for the Higgs boson at CERN. Competitions have resulted in successful projects including furthering HIV research and traffic forecasting. The learnings that result from the competitions and shared on Kaggle are being transferred into enterprise workflows to transform how business works.

Visit www.kaggle.com to learn more.



THE POWER OF Z:

1.5X faster speed of processing

Cut algorithm development time by at least 50%

Can try almost any experiment; no constraints on size

Preloaded software stack saves setup time

creating algorithms by approximately 50%. Wu elaborates, “I mainly used U-Net with EfficientNet as the encoder. With a 24GB RTX 6000 GPU, I could train the models with 640 x 640 images and a batch size of 16. My training set consisted of about 18,000 images, and an epoch took me about 17 or 18 minutes. If I trained the model with dual GPUs, one epoch would take less than 10 minutes. On the other hand, with my previous 11GB GPU, I could hardly train the model with a batch size of four. One epoch would take about 21 or 22 minutes.”

The future of data science

Data Science is still in its early stages, which is ideal for people wanting to work in this field. Job seekers will find there are several specialties to choose from if their goal is to provide value to a company or enrich people’s lives. It’s also strategic for enterprises looking to make their processes more efficient, people’s jobs less labor-intensive, and their companies more competitive.

Wu gives an example of how data science is helping him in his career with News Break.

“We have about 12 million users in the U.S., so that’s a lot,” he says. “Our operation needs good algorithms and good models. We need to classify our stories into different categories and geo-tag them for access by the right users. It’s very involved in a country as large as the U.S. with many cities that have the same name.”

It should be readily apparent to those with even a cursory knowledge of data science how the inherent methods, algorithms, and systems can be applied. It should also be clear that some companies, like News Break, can only stand to gain from the benefits of data science.

Wu predicts that developments in the field will continue to progress at an impressive clip. One possible path is to go deeper into Big Data. “But there will still be situations where we are working with small data too, where it’s difficult to collect samples. We cannot always collect enough data to work with, so there is still much work to be done.” And like the child who learned to program so he could draw on a computer, he’s fascinated by the possibilities.



LET US HELP YOU CREATE SOME AMAZING BUSINESS SOLUTIONS TODAY

CONTACT US

© Copyright 2021 HP Development Company, L.P. The information contained herein is subject to change without notice. The only warranties for HP products and services are set forth in the express warranty statements accompanying such products and services. Nothing herein should be construed as constituting an additional warranty. HP shall not be liable for technical or editorial errors or omissions contained herein.

As a Z by HP Global Data Science Ambassador, Yuanhao Wu's content is sponsored and he was provided with HP products.

4AAB-0310ENW, February 2022

