2020 AI & Data Science Silicon Valley, USA Online Internship Program

5 weeks programs : From July 13th ~ August 14th

Headstart Silicon Valley,
ASCAAI (American Society for Convergent Applications in AI),
We The People USA

Introduction

This program is a project-based program for college students who want to experience the culture of working in Silicon Valley and create a list of personal projects related to artificial intelligence and data science. Experts currently working in major IT companies and universities in Silicon Valley will provide students with the right skills and requirements to learn as future cutting-edge technology leaders.

Schedule & Curriculum

5 weeks programs : From July 13th ~ August 14th

1WK: Introduction and team building, pick team project, kickoff meeting

2WK: Mentoring day, team presentation, lecture 1 & 2, team review meeting 1

3WK: Mentoring day, team presentation, lecture 3 & 4, team review meeting 2

4WK: Mentoring day, team presentation, lecture 5 & 6, team review meeting 3

5WK: Mentoring day, make team presentation in English at **virtual online presentation at AI JAM -US (Aug 14th), www.aijam-us.com**

Daily Schedule

Mon	Thu	Wed	Thu	Fri
Day 1 (Jul 13)	**Day 2 (Jul 14)**	**Day 3 (Jul 15)**	**Day 4 (Jul 16)**	**Day 5 (Jul 17)**
• Opening Talk - Self Introduction, Details of Programs, Staff & Instructors • Introduction of Company & Program	• Group Interview 1 • Introduce group projects and requirements for students to choose	• Group Interview 2 • Introduce group projects and requirements for students to choose	• Group Interview 3 • Introduce group projects and requirements for students to choose • Team building 1	• Team building 2 • Work on **Internship Project** • SV Talk – Ryan Koo
Day 8 (Jul 20)	**Day 9 (Jul 21)**	**Day 10 (Jul 22)**	**Day 11 (Jul 23)**	**Day 12 (Jul 24)**
• Check-up meeting • Work on **Internship Project**	• Detail of projects and goal • Lecture 1 & Discussion Ho Joon Lee - Stanford Biomedical Data Scientist	• Work on **Internship Project** • Check-up meeting	• Work on **Internship Project** •Lecture 2 & Discussion Andrew Park - Palo Alto Network, Manager of Big Data Engineer	• Work on **Internship Project** • SV Talk – Eunyoung Kim, Sr. Software Engineer @ Google • Review project meeting by team Andrew Park, Ho Joon Lee
Day 15 (Jul 27)	**Day 16 (Jul 28)**	**Day 17 (Jul 29)**	**Day 18 (Jul 30)**	**Day 19 (Jul 31)**
• Check-up meeting •Work on **Internship Project**	•Lecture 3 & Discussion - Ho Joon Lee •Work on **Internship Project**	•Work on **Internship Project** • Check-up meeting • SV Talk – Driverless Driving Dr. Youngwook Kwon, Phantom AI	•Lecture 4 & Discussion - Ho Joon Lee, Andrew Park •Work on **Internship Project**	•Work on **Internship Project** • SV Talk – Startup • Review project meeting by team Andrew Park, Ho Joon Lee
Day 22 (Aug 3)	**Day 23 (Aug 4)**	**Day 24 (Aug 5)**	**Day 25 (Aug 6)**	**Day 26 (Aug 7)**
•Work on **Internship Project** • Check-up meeting	•Lecture 5 & Discussion - Ho Joon Lee, Andrew Park •Work on **Internship Project** •Review project meeting	•Work on **Internship Project** • Check-up meeting • SV Talk – Motion2AI, Byungsoo Kim, CEO	•Lecture 6 & Discussion - Ho Joon Lee, Andrew Park •Work on **Internship Project** • SV talk – Kee Yong Han, Udemy, CTO	• Online Public Presentation by Group in Korean
Day 29 (Aug 10)	**Day 30 (Aug 11)**	**Day 31 (Aug 12)**	**Day 32 (Aug 13)**	**Day 33 (Aug 14)**
•Work on **Internship Project** • Check-up meeting	•Lecture 7 & Discussion - Ho Joon Lee, Andrew Park •Work on **Internship Project** •Review projects meeting	•Discussion, Q&A •Work on **Internship Project** • Check-up meeting	•Team Project Presentation •Final Public Presentation in English - AI JAM - US Competition	•Closing team meeting and review

Instructors & Mentors

<u>Head of Instructor:</u>
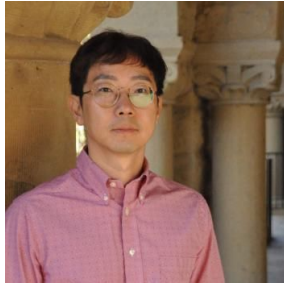
Sr Biomedical Data Scientist at Stanford University

Ph.D Ho Joon Lee,

- Senior Research Engineer, Stanford University (3/16/2019 - Present)
- Project Leader, Stanford University (3/16/2017 - 3/15/2019)
- Post-doctoral fellow, Stanford University (3/16/2012 - 3/15/2017)
- Research Associate, Arizona State University (8/2005 - 2/2012)

Manager Big Data Engineering, Palo Alto Network,
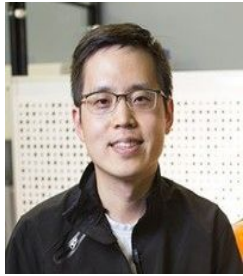
Andrew Park

- Former manager of Engineering, LinkedIn, Samsung

AI Autonomous Driving, Computer Vision + Deep Learning @Phantom AI,

PhD @ University of Berkeley

Dr. Youngwook Kwon

AI Mechanical Application,

CEO, Motion2AI,

AI for warehoue efficiency and safety

 CEO, Byungsoo Kim,

CTO, @ Udemy,

Former Director of Engineering @ Yahoo

Kee Yong Han

Data Scientist @Snap Inc,

Fomer Senior Applied Researcher @Microsoft

PhD Jin Young Kim

Sr. Software Engineer @ Google,

PhD @ University of Southern California

Eunyoung Kim

AI Mechanical Application,

Machine Learning expert @ Hacker Dojo

Dr. Mike Bowles

AI Application and Business,

Industry Fellow @ UC Berkeley

Rick Rasmussen

Benefits after the online internship program

- Certification of Internship
- AI JAM International Competition Awards (www.aijam-us.com)
- After the project, we offer qualified students the opportunity to take long-term internships to develop AI / data science courses at Flagly (www.flagly.org)

Video, interview with teams in 2019

http://y2u.be/mczqzoQPDBk

AI & Data Science Projects - title and detail

Precision medicine, also known as personalized medicine, by big data analysis of genomics

1. Cancer genomics for the optimal treatment

Mutations occur in the genome, which will alter the products of a much smaller number of genes involved in regulating cell growth, which then accumulates in developing cancer. Most often these mutations are unique to individual patients, highlighting the need for personalized approaches. Recent advancements in next-generation sequencing (NGS) technologies enable to sequence the genome of cancer patients and have created promising opportunities for precision medicine.

Open Questions:

- How to identify the genomic alterations that affect clinical outcomes such as drug

response, prognosis, and survival?
- How to determine the optimal treatment for each cancer patient based on its genomic profiling?

Tools used (what student will learn)

- Any coding skill (preferably Python) is required.

- Exposure to basic sequence analysis tool and web sites

- Basic understanding of cancer genomics analysis

- Basic understanding of cancer clinical outcomes

- Basic understanding of statistical analysis

2. The analysis of SARS-CoV-2 for preparing the COVID-19 pandemic

Background: The current sequencing technologies quickly generated multiple SARS-CoV-2 genome sequences. We have more >7,000 genomes of SASR-CoV-2 in 4 months since it's an outbreak. The analysis of these genomes provides insight into their origins, evolutionary trace, drug/vaccine targets. This information is crucial to handle pandemic successfully.

Open Questions:

- How to identify the difference between the two viral genomes?

- How to quantify the difference between the two genomes?

- How to characterize the SARS-CoV-2 genomes?

Tools used (what student will learn)

- Any coding skill (preferably Python) is required.

- Exposure to basic sequence analysis tool and web sites

- Basic understanding of viral genomics analysis

- Basic understanding of statistical analysis

Required Technical Skills:

Docker, Kubernetes, Python Frameworks, Java Frameworks, Mobile Application, Cloud, Python PyTorch, Google Tensorflow is preferred.

	Python	*Java*	*Javascript*	*Database*
Framework	Django Flask	Spring Play Hibernate(ORM) Maven, Gradle	jQuery React Angular Redux TypeScript	SQL MySQL Oracle Mongo

Tools	Pip Sublime IntelliJ	IntelliJ Eclipse	Node.js Phantom Selenium Jasmine	Workbench SQL clients

Open source projects in 2020

In addition, there are more optional projects, which students can upgrade the result of previous project, which have been open source to public in 2019 below:

- Style conversion project: A project to convert sketchy images of ordinary people into images of famous artists
- Machine learning project: A project to distinguish faces among groups using machine learning techniques

Samples of AI/Data Science Project in 2019

https://github.com/come-back-home/petective_web : Picture conversion from low resolution to high resolution
https://github.com/hsvc : Drawing style conversion project
https://github.com/onlyyou-teamb : Face recognition Project

인공지능 & 데이터과학 프로젝트 개요

유전체학 Big Data 분석을 통한 환자마다 최적화된 의약품 개발에 대한 프로젝트

본 프로젝트는 Stanford 약학/Biomedical 연구소에 계신 Sr. Big Data Scientist 이호준 박사님의 주도로 진행되는 프로젝트입니다. 수많은 데이터 중 "의료 및 신약개발관련 빅데이터를 다루는 기회를" 얻기 쉽지 않습니다. **본 프로젝트를 참여하는 학생들에겐 데이터과학를 다루는 인력 시장에서 매우 회귀하고 특별한 경력과 경쟁력을 갖출 수 있는 기회입니다**. 특히, 실리콘밸리에선 어떠한 방식으로 빅데이터를 분석하며, 인공지능을 어떠한 방식으로 활용하는지 미국 최고의 학교와 회사에서 종사하시는 분들에게 직접 배울 수 있는 소중한 기회가 될 것입니다. Biomedical 분야의 데이터 분석을 통해, 다양한 분야의 데이터에 적용한 응용 분석 능력을 갖추기를 희망하는 학생들에게 적극 추천합니다.

1. 최적화된 암치료를 위한 유전체학

돌연변이는 게놈에서 발생하며, 이는 세포 성장 조절에 관여하는 훨씬 더 적은 수의 유전자의 산물을 변화시켜 암 발생에 축적됩니다. 대부분 이러한 돌연변이는 개별 환자마다 고유하므로 개인화 된 접근 방식의 필요성을 강조합니다. 차세대 시퀀싱 (NGS) 기술의 최근 발전은 암 환자의 게놈을 시퀀싱하고 정밀 의학에 대한 유망한 기회를 창출하였습니다. 본 프로젝트를 통해, 데이터분석을 통한 치료약개발에 대한 과정이해와 방법을 경험해 볼 수 있습니다.

2. COVID-19 (코로나 바이러스) 유행성 전염병 대비를 위한 SARS-CoV-2 분석

현재 심각한 전염병으로 간주되는 코로나 바이러스에 대한 연구 프로젝트입니다. 현재 시퀀싱 기술은 여러 SARS-CoV-2 게놈 서열을 신속하게 생성했습니다.

코로나 사태가 발생한 이후, SASR-CoV-2가 7,000 개 이상 만들었으며 이들 게놈의 분석은 전염병의 기원, 진화 적 흔적, 약물 / 백신 표적에 대한 통찰력을 약품개발자들에게 제공합니다. 본 프로젝트를 통해, 게놈의 차이 식별하는 분석과 게놈의 특징을 발견하는 분석을 경험하게 될 것입니다.

Open source 활용/응용 프로젝트 in 2020

추가적으로 선택할 수 있는 프로젝트로 기존의 오픈소스를 활용하여 본인의 아이디어를 구현하거나, 응용 서비스를 만들어내는 프로젝트

- 스타일 변환 프로그램 : 평범한 사람들의 스케치 이미지를 유명한 예술가의 이미지로 변환하는 프로젝트를 변환 및 응용하여 프로젝트 수행
- 사물의 이미지를 구별기능 응용 프로그램 : 기계 학습 기술을 사용하여 다양한 객체들이 모여있는 사진에서 객체를 구별하여 특정의 서비스나 문제점을 해결해 주는 프로젝트

참고용 샘플들, AI/Data Science Projects in 2019

https://github.com/come-back-home/petective_web : Picture conversion from low resolution to high resolution
https://github.com/hsvc : Drawing style conversion project
https://github.com/onlyyou-teamb : Face recognition Project

Contact us:

Headstart Silicon Valley, Inc.
3350 Thomas Road, Suite 150
Santa Clara, CA 95054

ryan.koo@headstartsv.com
www.headstartsv.com