Isabel Chien

ADDRESS: 1 Holland St, Cambridge, UK CB4 3DL
PHONE: +44 (0)7951380796 | **EMAIL:** ischien3@gmail.com

Education

University of Cambridge
Ph.D. in Engineering, Machine Learning Group | Sept 2020 – Present
- Supervisors: Richard Turner (Cambridge), Aditya Nori (Microsoft Research)
- Funding via Microsoft Research PhD Scholarship Programme

Massachusetts Institute of Technology
M.Eng. in Computer Science and Engineering – GPA 5.0/5.0 | June 2018
- Advisor: Manolis Kellis
- Thesis: Natural Language Processing for Precision Clinical Diagnostics and Treatment (Grade: A+)

B.S. in Computer Science and Engineering – GPA 4.7/5.0 | June 2017

Skills

Fluent: Python, C#, Java, JavaScript, HTML/CSS
Proficient: R, Swift (iOS and watchOS), Scala, C++, Ruby (on Rails), Typescript, MATLAB

Publications

Alvin Shi, Gyulnara G Kasumova, William A Michaud, Jessica Cintolo-Gonzalez, Marta Díaz-Martínez, Jacqueline Ohmura, Arnav Mehta, **Isabel Chien,** Dennie T Frederick, Sonia Cohen, Deborah Plana, Douglas Johnson, Keith T Flaherty, Ryan J Sullivan, Manolis Kellis, Genevieve M Boland. "Plasma-derived extracellular vesicle analysis and deconvolution enable prediction and tracking of melanoma checkpoint blockade outcome." *Science Advances,* November 2020.

Richard E Leiter, Enrico Santus, Zhijing Jin, Katherine C Lee, Miryam Yusufov**, Isabel Chien**, Ashwin Ramaswamy, Edward T Moseley, Yujie Qian, Deborah Schrag, Charlotta Lindvall. "Deep Natural Language Processing to Identify Symptom Documentation in Clinical Notes for Patients With Heart Failure Undergoing Cardiac Resynchronization Therapy." *Journal of Pain and Symptom Management*, November 2020.

Isabel Chien#, Angel Enrique#, Jorge Palacios#, Tim Regan#, Dessie Keegan, David Carter, Sebastian Tschiatschek, Aditya Nori, Anja Thieme*, Derek Richards*, Gavin Doherty*, Danielle Belgrave*. "A machine learning approach to understanding patterns of engagement with internet-delivered mental health interventions." *JAMA Network Open,* July 2020.

Rebecca J Weiss, Sara V Bates, Ya'nan Song, Yue Zhang, Emily M Herzberg, Yih-Chieh Chen, Maryann Gong, **Isabel Chien**, Lily Zhang, Shawn N Murphy, Randy L Gollub, P Ellen Grant, Yangming Ou. "Mining multi-site clinical data to develop machine learning MRI biomarkers: application to neonatal hypoxic ischemic encephalopathy." *Journal of Translational Medicine*, December 2019.

Alex Chan*, **Isabel Chien***, Edward Moseley, Saad Salman, Sarah Kaminer Bourland, Daniela Lamas, Anne M Walling, James A Tulsky and Charlotta Lindvall. "Deep Learning Algorithms to Identify Documentation of Serious Illness Conversations During Intensive Care Unit Admissions." *Journal of Palliative Medicine*, November 2018.

Charlotta Lindvall, EJ Lilley, Sophia Zupanc, **Isabel Chien**, Brooks V Udelsman, Anne Walling, Z Cooper, James A Tulsky. "Natural Language Processing to Assess End-of-Life Quality Indicators in Cancer Patients Receiving Palliative Surgery." *Journal of Palliative Medicine,* October 2018.

Brooks Udelsman, **Isabel Chien**, Kei Ouchi, Kate Brizzi, James A Tulsky, and Charlotta Lindvall. "Needle in a Haystack: Natural Language Processing to Identify Serious Illness." *Journal of Palliative Medicine,* September 2018.

L. Zack Bright, Michael Handley**, Isabel Chien**, Sebastian Curi, L. Anders Brownworth, Sebastian D'hers, Ulrich R. Bernier, Pablo Gurman, Noel M. Elman. "Analytical models integrated with satellite images for optimized pest management." *Precision Agriculture*, 2016.

#, * denotes equal contributors

Preprints & Working Papers

Niranjani Prasad**, Isabel Chien**, Tim Regan, Angel Enrique, Jorge Palacios, Dessie Keegan, Usman Munir, Ryutaro Tanno, Hannah Murfet, Aditya Nori, Derek Richards, Gavin Doherty, Danielle Belgrave, Anja Thieme. "Deep Learning for the Prediction of Clinical Outcomes in Internet-Delivered CBT for Depression and Anxiety." Under review.

Umang Bhatt, **Isabel Chien**, Muhammad Bilal Zafar, Adrian Weller. "DIVINE: Diverse Influential Training Points for Data Visualization and Model Refinement." *arXiv.* Under review.

Workshops & Posters

Isabel Chien, Alvin Shi, Alex Chen, Charlotta Lindvall. "Identification of Serious Illness Conversations in Unstructured Clinical Notes Using Deep Neural Networks." *Joint Workshop on AI in Health (AIH), ICML, IJCAI, AAMAS* 2018, Stockholm, Sweden.

Isabel Chien, Alex Chan, Edward Moseley, Saad Salman, Sarah Kaminer Bourland, Daniela Lamas, Anne M Walling, James A Tulsky and Charlotta Lindvall. "Identification of Serious Illness Conversations During Intensive Care Unit Admissions Using Deep Neural Networks." Poster and spotlight presentation at the *2nd NorthEast Computational Health Summit (NECHS),* April 2018, Cambridge, Massachusetts.

Alvin Shi*, Gyulnara Kasumova*, **Isabel Chien**, Jessica Cintolo-Gonzalez, Dennie T. Frederick, Roman Alpatov, William A Michaud, Deborah Plana, Ryan Corcoran, Keith Flaherty, Ryan J Sullivan, Manolis Kellis, Genevieve M Boland. "Deconvolution of plasma-derived exosomes for tracking and prediction of immunotherapy across multiple tissues." Poster at *American Association for Cancer Research,* April 2018, Chicago, Illinois.

Isabel Chien, Lily Zhang, Randy L. Gollub, Rebecca J Weiss, Emily M Herzberg, Yih-Chieh Chen, KS Krishnamoorthy, Joseph H. Chou, Steve Pieper, Rudolph Pienaar, Lillia Zolei, Ellen Grant, Sara V Bates, Yangming Ou. "Machine Learning Detection of Hypoxic Ischemic Injury in Neonatal Brain MRI." Poster at *MGH (Massachusetts General Hospital) for Children Clinical Research Day*, March 2018, Boston, Massachusetts.

Alvin Shi*, Jessica Cintolo-Gonzalez*, **Isabel Chien**, Dennie T. Frederick, Roman Alpatov, William Michaud, Deborah Plana, David Panka, Ryan Corcoran, Keith Flaherty, Ryan Sullivan, Manolis Kellis, Genevieve M Boland. "Exosomal transcriptomic signatures tracks and predicts response to checkpoint blockade immunotherapy." Poster at *American Association for Cancer Research Special Conference on Tumor Immunology and Immunotherapy*, October 2017, Boston, Massachusetts.

Research Experience

Machine Learning Group – University of Cambridge
Supervised by Professor Richard Turner | Ph.D. Researcher | September 2020 – Present
I am interested in fairness and ethics in machine learning, particularly in forms relevant to healthcare or other "social good" applications. I am currently interested in characterizing fairness in exploration during learning to develop methods and measures of fair optimal decision making. In dynamic settings, such as bandit learning or online learning, the burden certain subgroups or individuals may bear in the tradeoff between exploration and exploitation could be considered an issue of fairness - what if a certain subset of the population incurs the cost of exploration at a disproportionate level? For this work, I am collaborating with clinical trials researchers to examine

the ideas of fairness in clinical trials (both randomized and adaptive). I am also working on interpretability for machine learning models through explanations by example.

Dana Farber Cancer Institute and Harvard Medical School, Department of Psychosocial Oncology and Palliative Care

Advised by Professor Charlotta Lindvall | Research Assistant | September 2017 – June 2018
Used natural language processing on unstructured clinical notes to extract palliative care information, such as patient and family care preferences, family meetings, and code status. Developed information extraction methods using recurrent neural networks and conditional random fields. To support this research, developed Python software (PyCCI) to annotate clinical notes with sensitive patient data. Also developed Python software (ClinicalRegex) to allow doctors and researchers to securely load clinical notes and run basic regular expression queries on them, typically as a baseline to more complex machine learning methods. PyCCI continues to be used by DFCI to conduct NLP research, and ClinincalRegex is currently being used to support clinical trials. Also used natural language processing to extract documentation of heart failure symptoms from unstructured clinical notes, using related machine learning methods. PyCCI annotation software was also used for this project. Posters and publications listed above.
PyCCI: https://github.com/chieni/PyCCI
ClinicalRegex: https://github.com/chieni/ClinicalRegex

Computational Biology Group - Computer Science and Artificial Intelligence Laboratory (CSAIL) at MIT

Advised by Professor Manolis Kellis | Master's and Undergraduate Researcher | September 2016 – June 2018
Analyzed exosomal signatures of melanoma patients to predict response to immunotherapy, in a collaborative effort with Dr. Genevieve Boland, a surgical oncologist at Massachusetts General Hospital, and her group. Worked with plasma-derived exosomal RNA samples from patients with metastatic melanoma undergoing anti-PD1 therapy from before treatment and during treatment. Aimed to discover potential biomarkers that stratify responders and non-responders prior to the start of immunotherapy and evaluate the performance of classifiers built using these biomarkers. This included creating and running statistical models and machine learning models and presenting and interpreting the results. Poster and publication listed above.

Martinos Center for Biomedical Imaging, Massachusetts General Hospital, Harvard Medical School & Department of Radiology, Boston Children's Hospital

Research Assistant | September 2017 – March 2018
Professor Randy Gollub, psychiatrist and professor at Harvard Medical School
Developed image analysis and machine learning algorithms for 3D MRI detection of brain abnormalities. Worked to label voxels of abnormal brain regions from MRI data for newborn babies with HIE (hypoxic ischemic encephalopathy) using computer vision techniques, including convolutional neural networks (investigated the application of U-Net). Also explored which families of traditional machine learning classifiers perform best on this type of task, on extracted MRI imaging features. Compared the labelled output of our algorithms with the overlap of our abnormal regions with multiple expert-annotated abnormal regions. Poster listed above.

Institute for Soldier Nanotechnologies – MIT

Undergraduate Researcher | Sept 2013 - Sept 2014
Designed and coded a dynamic interface for selecting and calculating an area that UAVs should cover, and launching those UAVs. Used HTML/CSS, JavaScript (jQuery), and Google Maps API. Aided in connection to the backend server with Node.js. Publication listed above.

Work Experience

Microsoft Research

Healthcare Intelligence Group – Research Engineer I & II | September 2018 – September 2020
Conducted applied machine learning research into a cognitive behavioral therapy platform. Constructed outcome prediction model for mental health patients using the platform and investigated methods of personalization and interpretation of user behavior. Responsible for exploratory machine learning work in probabilistic modelling and deep learning, helping determine the future direction of research, as well as helping write academic papers. Acted as a co-chair on the diversity and inclusion committee.

Verily (an Alphabet/Google subsidiary)
Diabetes Management - Software Engineering Intern | Summer 2017
Member of the iOS team of the Diabetes Management platform. Expanded the platform to Apple Watch by independently specifying, designing, and executing a completely new watchOS application. Also contributed to other features on the iOS application. Developed and wrote an internal guide to Apple Watch development. Quickly learned iOS and watchOS development with Swift.

Counsyl
Genomics - Software Engineering Intern | Summer 2016
Developed and implemented an algorithm that uses family history data to inform classification of genetic variants of unknown significance. Implemented a pipeline that runs once a day, retrieves family history data from consented Counsyl patients, runs it through the algorithm, and outputs assessed disease risk.

Tumblr
Services - Software Development Intern | Summer 2015
Architected and implemented a service to handle a high volume of requests to deliver push notifications to iOS and Android devices. Any time an event occurs that requires a push notification (such as liking, reblogging posts, new posts), a request to the service is made. Designed the service to handle at least 8,000 requests per second and wrote it in Scala, using Tumblr's open-source Colossus I/O framework and distributed systems principles. Unit tested and load tested the service to satisfaction, which is now rolled out to 100% of Tumblr users.

Thomson Reuters
Data Innovation Lab - Data Science Intern | January 2015
Contributed to the identification of 100,000 private companies deemed as significant (via various factors such as revenue, venture capital funding, ownership of patents) enough to be consistently researched. Performed web crawling and the use of APIs to scrape and analyze data from various sources.

Handy
Software Development Intern | Summer 2014
Executed full stack re-design of several web pages using HAML/SASS, AngularJS, CoffeeScript, jQuery, Ruby. Restructured Ruby on Rails code using AngularJS for cleaner and more efficient functionality. Created a Ruby API for dealing with displaying job distribution and connected it to the front-end. Built a web crawling feature to retrieve and document significant customer review data.

Philips Healthcare
AED Research - Research Intern | September 2012 - September 2013
Developed and created a MATLAB application for reading, annotating, and syncing EKG waves. Completed data analysis for accelerometer and EKG readings. Refactored existing MATLAB functions to improve usability. Completed data analysis for EKG wave readings used for clinical research using MATLAB, MOW, and MMCDataReader. Redacted documents related to clinical trials.

Microsoft
Microsoft Research, Translate Team - Research Intern | Summer 2012
Worked in a three-person team to create a web-based multilingual, multiplayer crossword game [crossword.microsofttranslator.com]. Responsible for the design and implementation of the user interface. Headed data gathering for the words and definitions lists of the thirty-five languages supported, working with text file manipulation, and using the Microsoft Translate API.

Teaching Experience

Introduction to Computer Science, Programming, Data Science (6.00)
Teaching Assistant - MIT | Fall 2017
Introductory programming course for over 500 students. Held office hours, managed lab assistants, graded technical assignments.

Computer Systems Engineering (6.033)

Teaching Assistant – MIT | Spring 2017
6.033 is a required course for all Computer Science majors at MIT that provides thorough instruction in computer systems. Taught select tutorials, graded technical papers and technical assignments, assisted in recitations, and held office hours.

Fundamentals of Programming (6.009)
Lab Assistant – MIT | Fall 2016
Held office hours, helping students with their programming lab assignments (Python projects).

Elements of Software Construction (6.005)
Lab Assistant – MIT | Spring 2015
Held office hours, helping students with their problem sets (Java programming assignments).

Honors and Awards

Microsoft Research PhD Scholarship Programme, 2020 - Present
Best Use of the AthenaHealth API, Grand Hack Healthcare Hackathon, MIT - 2017
Scholar, MIT Arts Scholars - 2017
Best Overall, Final Project Award for Software Studio (6.170), course with over 300 students, MIT - 2016
National Merit Scholar, 2013-2017
National AP Scholar, 2012