		STEVEN YAN
Email:	datascisteven@gmail.com	
Portfolio:	datascisteven.github.io	
Mobile:	646.991.4222	
Connect:	@datascisteven @datascisteven @datascisteven	

TECHNICAL SKILLS

Machine Learning (*Scikit-learn, Tensorflow***), SQL (***MySQL***), Python (***NumPy, Pandas, SciPy***), Data Analysis, Data Visualization (***Matplotlib, Seaborn, Tableau***), Git, Microsoft Office, HTML/CSS**

TECHNICAL PROJECTS

Melanoma Image Classifier for Mobile App Deployment – *Github*
- Developed CNN model from scratch alongside pretrained models to detect melanoma trained from 33,126 images in ISIC 2020 dataset ready for deployment for app accessible to all populations
- Evaluated performance on two holdout sets achieving 95%/80% accuracy with VGG19 after addressing class imbalance by bolstering with additional datasets and data augmentation with Albumentations

Automated Hate Tweet Detection – *Github*
- Developed NLP binary classification model for detecting hate tweets downloaded from Twitter API to address growing concern of hate speech to incite recent hate crime
- Implemented a SVM algorithm with minority class addition to achieve holdout accuracy of 92.0%, with an eye to two other metrics: high ROC-AUC of 0.965 and lowest FN%'S at 6.21%

Credit Card Default Prediction – *Github*
- Developed classification model to predict whether customer will default on next payment as well as identifying main factors for customer defaulting trained on dataset of 30000 customers
- Implemented resampling techniques from imbalance-learn and smote-variants across different models and engineered novel features for classifying CC default achieving 82% accuracy

RECENT EXPERIENCE

Part-Time Volunteer, DataKind   **Current**
DataKind DataDive Event [Github], *Exploring Equity in Broadband Access in US (UChicago CDAC)*   **9/2021**
Flatiron Hackathon [Github], *SE, DS, and UIUX Alumni Collaborate to Build a Gaming App*   **8/2021**

WORK EXPERIENCE

Educational Consultant (Freelance), *Examsherpa, New York, NY*   **8/2009–10/2020**
- Developed individualized study plans for premeds based on assessment and academic history
- Projects included: developing and delivering 20-session bootcamp for students in Mentoring In Medicine, a pipeline program for underrepresented premeds, collaborating with content writers led by Khan Academy and AAMC (test-making organization) to develop open-source high-quality test prep materials in effort to equalize access to resources for all

Curriculum Developer/Teacher Trainer, *Kaplan Test Prep, New York, NY*   **2/2006-5/2009**
- Wrote and developed MCAT Tutoring Guide for new tutors based on my training/teaching experience
- Managed team of 15 instructors in Manhattan/Bronx region by assigning classes, conducting class observations, and providing feedback to increase student excellence ratings for region
- Conducted regular prospective instructor auditions and training course for new hires

EDUCATION

Flatiron School, *Data Science Online Immersive Program, NYC Campus*   **10/2020–4/2021**
University of Chicago, *BA Biology, 3.7 GPA, College Honor Scholarship*   **9/1996–6/2000**