# kacey Wang

SOFTWARE DEVELOPER / APPLIED SCIENTIST

Los Angeles, CA   323-337-6337
wang.works/   kacey@ieee.org
kaceywang   kaixi-wang

## SUMMARY OF QUALIFICATIONS

Creative, empathetic, self-driven developer who uses data to develop *intelligent* solutions that are *intuitive* to users and *explainable* to collaborators

> Leverages strengths in adaptability, lateral thinking, and effective communication to quickly integrate into and begin contributing on existing development teams

> Fluent in using Tensorflow and PyTorch to build deep learning models in Python, with experience in both training models from scratch and fine-tuning pre-trained models for domain-specific tasks

> Skilled at mining high-quality information from heterogeneous data sources and delivering reliable insights by judiciously implementing data-driven machine learning algorithms and domain-driven rule-based methods

### TECHNICAL SKILLS & TOOLS

- **Python**, JavaScript, Java, MATLAB

  matplotlib, numpy, pandas, pyspark, scipy, sklearn// tensorflow, keras, pytorch//gensim, fasttext, spacy, textblob, nltk

- Algorithms, Data Structures, Big-O
- Hadoop/Spark MapReduce, Dataproc
- Database systems, SQL, noSQL
- Multimedia systems, signal processing

- PHP, Node.js, HTML, CSS
- Git, GCP, RESTful APIs
- EN (native), ZH (fluent), ES (6yrs)

Areas of Expertise **Machine Learning/Deep Learning, Data Analysis, Information Retrieval, NLP, Web Development**

## EDUCATION

**M.S. COMPUTER SCIENCE | UNIVERSITY OF SOUTHERN CALIFORNIA** (Los Angeles, CA)    Jan 2018-Dec 2019

Courses: Deep Learning, Machine Learning, Foundations of Artificial Intelligence, Analysis of Algorithms, Information Retrieval & Web Search Engines, Web Technologies, Database Systems, Multimedia Systems Design

**B.S. NEUROSCIENCE, BIOMEDICAL ENGINEERING | TULANE UNIVERSITY** (New Orleans, LA)   Aug 2013-May 2015
*magna cum laude*

## WORK EXPERIENCE

**TEACHING ASSISTANT | USC DEPARTMENT OF COMPUTER SCIENCE**    Jan 2019-Dec 2019

> Co-produced the graduate-level, 200+ student course CSCI-572: Information Retrieval & Web Search Engines

> Topics: web crawling; NLP, document parsing, data indexing, query processing; big data ecosystems (Hadoop, MapReduce); sorting and ranking algorithms; Google business logic, SEO, Adsense, Ad Exchange

**PROJECT MANAGER, RESEARCH SCIENTIST | BJC INSTITUTE OF HEALTH**    Jun 2015–Dec 2017

> Designed "neuron-on-a-chip" devices for modeling and analyzing complex neural networks in vitro

> Optimized methods for rapidly fabricating microfluidic devices, achieving feature resolutions down to 10 microns without specialized equipment

> Mentored interns in designing research experiments, maintaining documentation, and analyzing data

**INTERN, ASSOCIATE RESEARCHER | WASHINGTON UNIVERSITY IN ST. LOUIS**    Jun 2010-Jun 2015

**SERVER | MULTIPLE CASUAL & FINE DINING RESTAURANTS (ST LOUIS)**    Jun 2010-Aug 2013

> Thrived working front-of-house for 40-80 hrs/wk, juggling multiple orders and customer requests simultaneously and generating $500-$1000+ in sales per 4-6 hr shift

> Received the *"Above and Beyond"* award at the Cheesecake Factory for outstanding guest feedback

# *PROJECTS*

### DOMAIN-SPECIFIC TEXT GENERATOR ⬀ | Deep Generative Models & Transfer Learning in NLP

> Built a Generative Language Model from scratch with PyTorch that outputs coherent long-form product reviews in response to inputted product keywords and ratings

> Achieved 49% improvement and alleviated posterior collapse by: (1) incorporating a conditional decoder and hyperparameter-tuned discriminator into the architecture of the Variational AutoEncoder (VAE), and (2) prioritizing data cleanliness

> Developed a comprehensive data wrangling strategy using PySpark; steps included profiling the metadata and analyzing the text using pre-trained sentiment analysis models to identify poor-quality or fake reviews in a semi-automated manner

> Presented a more promising solution in supplementary work using transfer learning. Generated realistic reviews with greater semantic and syntactic variety after just 30min of training by using OpenAI's massive pre-trained model (GPT-2 355M) in place of an untrained model

### DYNAMIC AD INJECTION ⬀ | Algorithm-based Feature Detection and Audio Signal Processing

> Developed a media player that not only transcodes raw digital media files, but also algorithmically replaces any existing commercials based on a brand's presence in the noncommercial content

> Implemented a fully affine invariant image comparison algorithm (A-SIFT) using OpenCV to match flat graphics to logos appearing in real-world scenes using OpenCV

> Reduced runtime >450% over baseline by: (1) refining rule-based heuristics to minimize the total number of video frames selected for feature matching without increasing the false negative rate, (2) analyzing the audio signal in the log-power domain to empirically validate an energy drop threshold that precisely identified scene boundaries

### SEARCH THE LA TIMES! ⬀ | Full-text Document Search Engine + Flask App Interface

> Built a news article search engine app powered by Apache Solr

> Created a Python-based client API/web UI using Flask that issued HTTP GET requests to the search backend and returned a Google-like search engine results page (SERP)

> Configured multithreaded Java web crawlers (crawler4j) to fetched  20,000 unique news articles from the LA Times for populating the search database

> Simplified mapper and reducer classes and dramatically reduced backend development costs by refactoring Hadoop MapReduce jobs into Spark streaming jobs

> App features full-text search, auto-correct based on Norvig's spell correction algorithm, data-driven phrase auto-suggest optimized with n-gram tries, optional PageRank scoring (calculated independently using NetworkX)

### WEATHER APP ×3 ⬀ | RESTful APIs in PHP, JavaScript (Angular + Node.js), and Swift

> Developed and presented a trio of weather apps built end-to-end using modern web frameworks

> Created a shared Node.js REST microservice which connects to external web APIs (Dark Sky and Google) to deliver real-time data securely to both web and mobile/iOS interfaces

> App Features: real-time weather forecasts for inputted addresses or current geolocation; multiple views rendered on demand with interactive charts and widgets for effective data visualizations; saving locations to local storage; social media integrations

### WANG.WORKS ⬀ | Django Web App

> Created a website in Python using Django to show samples of my work and serve as a playground for learning about or experimenting with design patterns and new web frameworks

> Deployed and managed with git to Heroku; Backend uses a Postgres database for storing unique page content and Google storage buckets to host media and static assets

---

## *AWARDS & HONORS*

> Future Leader in Tech  Women Impact Tech Conf. LA 2019

> Research protocols published in peer-reviewed journals

## *HOBBIES & INTERESTS*

> Learning, DIY home decor and upcycling projects, listening to podcasts, drawing & painting, fitness, beauty & style