

SUMMARY

Data scientist specialized in GIS/Spatial data mining/Data analytics/Visualization/Energy/Business/

SKILLS & EXPERTISE

- Language: **Python** (pandas, geopandas, matplotlib, shapely), SQL, R
- Machine learning (sklearn): regression, classification, cluster
- **GIS: ArcGIS, QGIS modeling**, Mapbox studio
- Database: Microsoft SQL, Crate, PostgreSQL.
- Platform: Linux, Ubuntu, Windows, IOS, Github
- Web: html, JavaScripts, OpenLayers

PROFESSIONAL EXPERIENCE

Picarro Inc.

San Francisco Bay Area, CA
2016 - present

Data Scientist

- Designed and developed the software solutions to improve product location accuracy using geospatial API, increasing the feature's market price by 78%;
- Implemented geoserver, docker and PostgreSQL, improving the product speed by 82%.
- Created innovative feature engineering approach of using spatial autocorrelation to resize feature to improve risk ranking model, helped the customer saved more than \$2.1M;
- Process, visualize and analyze large volume of data using Python, QGIS, SQL on a daily basis.
- Integrated clustering and analyzed hazardous gas leaking risk supporting marketing team to win the more than \$10M contract.

Visitor

- Published paper on using genetic algorithm to optimize the gas emission characters.
- Published paper on the impacts of geographic coordinate system on the numeric model
- Developed open source code to read and visualize a complex scientific data NetCDF in GIS.

Research Assistant

- Used regression, neural network, K-mean, tree to predict greenhouse gas leaks.
- Collected 600 GB geographic data; Used GIS modeling to estimate methane emissions from variety of sources; Created web application using HTML and JavaScript and deployed Geoserver and Thredds server to display GIS layers, NetCDF and GeoJSON files
- Published research articles, interviewed by media, and presented at top conferences

Intern

- Bank lobby manager assistant
- Helped customers fill out the forms, recommended financial products

Data Analyst Intern

- Dedicated to Beijing urban planning Road Name Database program, performed spatial analysis, create thematic maps; Consulted 3D Beijing urban planning system conference

**NATIONAL CENTER FOR
ATOMOSPHERIC RESEARCH
(NCAR)**

Summer 2015

**THE PENNSYLVANIA STATE
UNIVERSITY**

Geoinformatics and Earth
Observation Laboratory
2014

**AGRICULTURAL BANK
OF CHINA (ABC)**

2012

**PIANNING AND
INFORMATION CENTER**

2014

EDUCATION

**THE PENNSYLVANIA STATE
UNIVERSITY**

University Park, PA, May 2016

Master of Science

- **Major:** Geography focusing on Geographic Information Science(GIS)
- **GPA:** 3.83 / 4.00
- Summer school at **Harvard University** (2013) and **Peking University** (2014)

AWARDS

- 2015 **Full Scholarship (\$42,415 per year)**
- 2009 **Second Prize in National High School Mathematic Olympic Competition**

LEADERSHIP

**INNOVATION &
INTERDISCIPLINARY
COMMUNITY (IICC)**

University Park, PA, 2015

PRESIDENT

- Facilitate interdisciplinary communication on academic topics, to promote critical and creative thinking, and support innovative projects in science and technology
- Organized more than 20 monthly interdisciplinary workshops

PERSONAL

- Fluent in English and Chinese
- Integrity
- Team player
- Tango learner
- Love Nutella
- World Traveler: traveled to 18 counties and celebrated my 23rd birthday in Arctic Circle.