

# John Grey

## Senior Data Scientist

Senior Data Scientist with 6 years experience who Collaborated with machine learning engineers and NLP engineers to design, develop, improve some Machine Learning Models on large datasets. Highly experienced in big data analysis and visualization. Possessed strong cooperation and communication.



jamesantontop@gmail.com



+1 8126182967



California, United States

## SKILLS

Statistics

Probability

C++

C#

R

machine learning

Python

Pandas

Matplotlib

Numpy

PyTorch

scikit-learn

Matlab

RDBMS

Big Data

spark

Hadoop

SQL

PostgreSQL

SQLite

NoSQL

Cloud

Tableau

Excel

## LANGUAGES

English

Full Professional Proficiency

Chinese

Professional Working Proficiency

## WORK EXPERIENCE

### Senior Data Scientist Sherpa.ai

05/2021 - 09/2022

*Achievements/Tasks*

- Developed Lung Diagnosis System using Iris Recognition Collaborating with machine learning engineers and Computer Vision engineers that accuracy is up to 98.5% and saved 10K+ persons so far
- Partnered with technical team to build out Fake Currency Detection System that detected 97.5% fake currency
- Upgraded and analyzed company's BI database that reduced size of company database to 72% and saved annual investment for about 150K

### Data Scientist Ecom Analytics Co

02/2018 - 01/2021

*Achievements/Tasks*

- Cleaned and integrated large sets of structured and unstructured datasets from the ground up for use in ML models and products that reduced size of dataset to 87% and improved model accuracy to 94.5%
- Worked with BI team and stakeholders to analyze and visualize their marketing data resulting in \$400K in annual incremental revenue
- Analyzed and visualized airline data with data analysis team improving airport traffic management system

### Data Scientist VIDEN - Data & Result-Driven Advertising and Digital Analytics Partner

08/2016 - 11/2017

*Achievements/Tasks*

- Built out marketing data infrastructure from the ground up using SQL to provide real-time insights into product
- Created and prepped data from NET to help data scientists build marketing mix models that lifted ROI by 4 basis points
- Collaborated with SEO team members to built out data infrastructure and analyzed saving 15% of annual investment

## PERSONAL PROJECTS

### Pneumonia Diagnosis System using Iris recognition (07/2021 - 07/2022)

- Prepared and parsed Iris dataset to improve accuracy 98.5% and provided insights of machine learning model developing

### Fake Currency Detection System (07/2020 - 11/2020)

- Using only 10K dataset, achieved over 99.87% accuracy on train dataset and 97.5% accuracy

### Resume Parse System (04/2020 - 06/2020)

- Parsed and visualized resume of applicants so that ML engineers select more accurate model

## EDUCATION

### Bachelor of Mathematics University of Bremen

09/2012 - 04/2016