Petros Venieris

pgvenieris@outlook.com +306946740374 Greece,Athens linkedin.com/in/petrosvenieris

github.com/venie1

Profile

Data scientist with proven expertise in predictive modeling, machine learning, and AI applications. Graduate from Georgia Tech master of science in analytics (3.91 GPA) with hands-on experience at Sandia Labs developing predictive maintenance models.

INTERACTIVE CV EXPERIENCE: Skip the usual resume!(For windows and android)
Experience my background through
my AI-powered assistant built with OpenAI GPT-4, advanced NLP, and prompt engineering:
https://resumeagent-5lxhhftrdnxwxgeys3dg2h.streamlit.app/

Education

MSc, Analytics, *Georgia Institute of Technology , GPA: 3.91/4.0* 12/2023 – 08/2025

- Top 5% of graduates in one of the best data science graduate programs in the world.
- Parnership with Sandia Labs for practicum.
- Developed end-to-end ML pipelines using Python (Pandas, NumPy, Scikit-learn), R (dplyr, ggplot2), SQL for predictive modeling, time series forecasting, SVM, Random Forest, AdaBoost, spectral clustering, Scikit-learn and neural networks
- Built algorithmic trading strategies using Python, Q-Learning, reinforcement learning, SciPy optimization.
- Developed Monte Carlo simulation models using Python, R, Arena Software for options pricing, discrete-event simulation, and stochastic optimization.
- Applied advanced regression, time series analysis, causal inference, A/B testing using R for financial risk assessment and marketing optimization
- Developed real estate analytics platform using Python, Facebook Prophet, XGBoost, Power BI.
- Led regression analysis of US housing prices using multiple approaches: macroeconomic factors, socio-economic variables, and market dynamics, achieving 88% explained variance in final ensemble model through advanced statistical techniques including BoxCox transformations, ridge regression, and logistic classification

BS, Digital Systems, *University of Piraeus(7,4 /10)* 09/2018 – 09/2022
Greece, Piraeus

- Thesis: COVID-19 forecasting Flask application (Grade: 10/10)
- Developed ETL workflows and deployed web applications.

Work Experience

Practicum: Data Scientist - Predictive Maintenance, *Sandia Labs* 05/2025 – 08/2025

- In talks with Sandia and Georgia tech Director for further investigation and research paper
- Ingested and merged over 1.1 million time-step records from 23,550 trucks; applied rigorous cleaning.
- Benchmarked Logistic Regression, Random Forest, Gradient Boosting, MLP and XGBoost on technical, behavioral and combined feature sets; achieved up to **0.732 ROC-AUC**, leveraged cost-sensitive thresholding to reduce average maintenance cost by **20–30%**, and validated with precision-recall analysis for imbalanced failures.
- Demonstrated that behavioral metrics can flag risky driving patterns early enabling targeted coaching to prevent failures while hybrid models optimize maintenance scheduling; scalable framework supports real-time fleet telematics and drives measurable safety and cost improvements.

- More information can be found here: https://github.com/venie1/Predictive-Maintenance-Behavioral-vs.-Technical-Analysis-of-Truck-Failures

Sales Associate, *Intersport* 01/2021 – 01/2023
- Balanced part-time work with full-time university coursework.

Sports Data Scientist, *BetMax, Self-Employed* 01/2019 – 12/2022
Greece, Athens
- Co-developed an automated live-football match prediction system with a partner.
- Processed real-time and historical data every 15 minutes, achieving 90% accuracy across 1,000+ matches.
- Automated stakeholder notifications via high-confidence email alerts, enabling consistent simulated ROI.
- More details can be found here https://github.com/venie1/BetPredictor

Projects

AI Resume Agent Project
- LLM application using OpenAI GPT-4 with sophisticated prompt engineering
- Engineered ML-powered data extraction pipeline processing structured/unstructured data (PDF, HTML, JSON) with automated feature engineering and skills classification algorithms
- conversational AI system with context-aware memory management and advanced NLP techniques.
- Developed responsive web interface using Streamlit, custom HTML/CSS.
- Technologies: OpenAI GPT-4, Python, Machine Learning, NLP, Streamlit, HTML/CSS, Plotly, Pandas, API
- more details can be found here: https://github.com/venie1/ResumeAgent

RealEstateAdvisor(from data collection to visual deployment, 6 people team job that lased 4 months)
- Built end-to-end ETL pipelines merging Redfin, FRED, and Census data sets.
- Engineered lagged and rolling features; trained RidgeCV, Random Forest, and XGBoost models (1-/3-month MAPE: -6%) and Prophet (6-/12-month MAPE: 7-10%).
- Deployed a Power BI dashboard with dynamic choropleth maps and synchronized filters; automated 80% of preprocessing and demonstrated a 12-15% simulated portfolio-return uplift as a ready to use dashboard.
- More details can be found here https://github.com/venie1/RealEstateAdvisor

Machine Learning Trading Strategies - Algorithmic Trading Portfolio(Class project)
- Built comprehensive algorithmic trading system using Python, ensemble ML models (BagLearner, Random Trees), Q-Learning, benchmark with 2.0+ Sharpe ratio on old dataset
- Implemented advanced technical analysis library (Bollinger Bands, MACD, RSI) with realistic market simulation including transaction costs and market impact modeling
- more details can be found here https://github.com/venie1/machine-learning-trading-strategies

COVID-19 Forecasting & Insights Application(thesis with 10/10 grade)
- Developed Flask web app with time-series forecasting, reducing RMSE by 15% from baseline
- Implemented interactive Plotly visualizations for spatial-temporal analysis
- More details can be found here https://github.com/venie1/Covid-prediction-application

Skills

Technical skills
- Programming: Python (Expert), R, SQL, Git, Docker, Flask, Streamlit
- Machine Learning: Scikit-learn, XGBoost, Random Forest, Neural Networks, PyTorch, TensorFlow
- AI/LLM: OpenAI API, GPT-4, Prompt Engineering, NLP, Natural Language Processing, Conversational AI
- Data Tools: Pandas, NumPy, Matplotlib, Plotly, Power BI, Jupyter

Specialized skills
- Time Series Forecasting: Prophet, ARIMA, SARIMA, Seasonal Decomposition
- Statistical Analysis: A/B Testing, Hypothesis Testing, Causal Inference, Monte Carlo Simulation
- Deep Learning: Computer Vision, Reinforcement Learning, Q-Learning, Feature Engineering