Project Summaries

Equity Copilot

title: "Equity Copilot | Dean Shabi", description: "RAG-powered AI assistant for the Equity research market, built on Azure AI and LangChain."
}; Project Overview Equity Copilot is an AI-powered platform that transforms how analysts and fund managers conduct investment research. By leveraging large language models and domain-specific fine-tuning, the system extracts nuanced insights from earnings calls, SEC filings, market data, and news sources. Traditional equity research is labor-intensive and time-consuming, with analysts spending countless hours reading transcripts, scanning financial statements, and cross-referencing data points. Equity Copilot automates much of this process, allowing human analysts to focus on higher-value activities like developing investment theses and strategic decision-making. Technical Solution Our approach combined several advanced AI technologies to create a comprehensive research assistant: Domain-Specific LLMs Fine-tuned foundation models to understand financial terminology, market dynamics, and accounting principles. RAG System Retrieval-augmented generation system that accesses proprietary financial databases and research reports to provide evidence-based analysis. Financial NLP Pipeline Custom natural language processing pipeline optimized for financial text, including sentiment analysis and named entity recognition for companies, executives, and key metrics. Multi-Agent Framework System architecture featuring specialized agents for different tasks, working together to analyze financial data and generate insights. System Architecture Data Sources SEC Filings Earnings Calls News & Reports Processing Pipeline NLP Engine Vectorization Data Enrichment Knowledge Base Vector DB Financial Metrics Market Data Agent Orchestration Layer Analyst Interface Reporting Engine Key Capabilities Earnings Call Analysis Extracting key insights, tonality shifts, and executive sentiment from quarterly earnings calls, highlighting significant changes from previous quarters. SEC Filing Summarization Automatically processing complex SEC filings (10-K, 10-Q, 8-K) to identify material changes, risk factors, and financial trends. Competitive Intelligence Tracking competitor mentions across earnings calls and public statements to map competitive dynamics and market positioning. Red Flag Detection Identifying accounting irregularities, concerning management changes, or operational challenges that might affect financial performance. Thematic Research Exploring emerging industry trends and themes by analyzing language patterns across multiple companies and sectors. Event-Driven Alerts Monitoring news, regulatory announcements, and social media for significant events that might impact investments. Business Impact 1 Research Efficiency Reduced initial research time by 60%, allowing analysts to cover 3x more companies with the same resources. 2 Enhanced Coverage Enabled comprehensive analysis of small and mid-cap stocks that previously received limited attention. 3 Insight Discovery Identified subtle patterns and connections across sectors that human analysts had overlooked. 4 Risk Reduction Earlier identification of company-specific risks by analyzing language changes in disclosures.);
}

Exempt Supply Matching

title: "Exempt Supply Matching | Dean Shabi", description: "Matching exempt supply and demand using neural networks and graph optimization algorithms."
}; Project Overview Created a unique matching algorithm operating under the UK's Supplier Exempt Class A and BSC Modification P442 regulations. This innovative solution optimally pairs SME energy consumers with local sub–5 MW generators, unlocking approximately £50/MWh in non-commodity cost savings. This algorithm has generated millions in new revenue streams and savings for businesses, while promoting more sustainable, localized energy consumption patterns. Benefits £50/MWh Cost Savings Up to 90% reduction in non-commodity costs £3M+ Value Generated 75% improvement over traditional solutions 200+ Successful Pairings 85% match success rate achieved Annual Savings Calculation 7 GWh Annual Generation × £50/MWh Cost Savings £350,000 Total Annual Benefit The Challenge UK renewable energy regulations offer significant cost-saving opportunities through "exempt supply" arrangements, but establishing these partnerships presents complex challenges: Regulatory Complexity UK energy regulations permit exemptions from certain non-commodity costs when generators supply nearby businesses directly, but navigating these regulations requires specialized expertise and careful compliance management. Matching Difficulty Finding viable generator-consumer pairs requires analyzing multiple complex factors: compatibility criteria, load profiles, suitable connection points, and technical feasibility. Scale & Efficiency Manually identifying and evaluating potential matches across thousands of sites is prohibitively time-consuming and prone to missed opportunities. Data Integration Combining and analyzing fragmented data from generation profiles, consumption records, grid infrastructure, and regulatory requirements presents significant technical challenges. Solution I developed a comprehensive solution to address the complex challenge of matching exempt renewable generators with nearby businesses, creating efficient and cost-effective energy partnerships that leverage UK electricity regulations. Technical Approach Advanced matching algorithm to identify optimal generator-consumer pairings Real-time regulatory compliance verification system Advanced load profiling to match generation and consumption patterns Secure blockchain-based contract management Key Innovations Proprietary scoring algorithm for optimal matching Dynamic regulatory compliance engine with real-time updates AI-powered consumption forecasting for maximizing exemption value Automated contractual agreement generation with legal validation Regulatory Framework Supplier Exempt Class A Regulatory classification that allows for certain exemptions from standard energy supply obligations when specific conditions are met between generators and consumers. BSC Modification P442 Balancing and Settlement Code modification that enables specific matching arrangements between small-scale generators and consumers, supporting localized energy markets. Key Regulatory Requirements • Generators must be sub-5 MW capacity to qualify for exemptions • Supply must meet regulatory requirements for direct supply • Matching must be documented and reported to regulatory authorities • Balancing responsibilities must be properly assigned and managed System Architecture Data Inputs Consumer Profiles Generator Output Location Data Processing Layer Matching Algorithm Optimization Engine Forecast Models Output Systems Match Reports Regulatory Docs Billing Integration Continuous Optimization Loop Application Scenario Example Pairing Solar Installation ~4.8 MW capacity solar farm with 7 GWh annual generation Business Complex A collection of 20-25 SMEs with varied energy needs Cost Calculation Annual generation: 7 GWh × £50/MWh savings = £350,000 potential annual benefit Key Benefits ' Cost Reduction Annual savings of £350,000 based on 7 GWh generation ' Efficient Energy Use Up to 85% of generated power consumed locally ' Revenue Stability More stable revenue streams for renewable generators ' Environmental Impact Carbon footprint reduction

equivalent to removing 150-200 cars from roads Business Impact 1 Cost Savings Unlocked approximately £50/MWh in non-commodity cost savings for participating businesses. 2 Revenue Generation Generated millions in new revenue streams through this innovative matching service. 3 Sustainability Promoted more sustainable, localized energy consumption patterns, reducing transmission losses. 4 Market Advantage Provided significant competitive advantage in the energy supply market with this unique offering. Future Developments Platform Scaling Expanding the platform to handle larger volumes of participants and more complex matching scenarios. Enhanced AI Implementing more advanced machine learning algorithms to improve matching efficiency and forecast accuracy. Marketplace Expansion Developing a broader marketplace model that supports additional energy services and participant types.);
}

Forecasting Models

title: "High-Accuracy Forecasting Models | Dean Shabi", description: "Developing advanced forecasting models for energy consumption, generation, and pricing."
}; Project Overview Led the development of long-term load, generation, and price forecasting models for energy-tech companies in the UK. These advanced models are now trusted by thousands of customers and have significantly enhanced accuracy, profitability, and market competitiveness. The forecasting systems include specialized components for PV generation, battery degradation, and price forecasting, all working together to power core product features, support trading decisions, and reduce balancing costs by providing accurate predictions at sub-hourly granularity. Technical Approach This project involved developing multiple specialized time series forecasting models: Load Forecasting Predicting consumption patterns across different customer segments with high granularity and accuracy. PV Generation Weather-aware forecasting that accounts for variables, panel degradation, and installation specifics. Battery Degradation Advanced models to predict performance decline over time for more accurate energy planning. Price Forecasting Sub-hourly granularity predictions to support trading and balancing decisions in real-time. The models integrate various data sources including historical consumption, weather forecasts, market data, and asset-specific parameters to deliver highly accurate predictions. Model Architecture Data Processing Pipeline Weather Data %¼ Historical Energy Data %¼ Market Signals %¼ ML Prediction Engine %¼ API & Integration Layer Business Impact 1 Core Product Features Powers critical functionality in energy management platforms used by thousands of customers. 2 Trading Decisions Supports critical trading strategies with high-confidence forecasts for optimal energy trading. 3 Cost Reduction Reduces balancing costs through precise sub-hourly predictions, saving millions annually. 4 Market Competitiveness Enhances overall positioning for energy suppliers through superior forecasting technology.);
}

Fruit Waste

title: "Reducing Fruit Waste with ML | Dean Shabi", description: "Computer vision models to detect fruit ripeness and reduce waste in the food supply chain."
};);
}

Mlops Tools

title: "Open Source MLOps Tools | Dean Shabi", description: "Comparing open source deployment and serving tools for machine learning models."
}; Project Overview The world of machine learning (ML) is as dynamic and diverse as it is complex, with numerous tools and practices aimed at streamlining and enhancing the deployment and serving of ML models. The challenge lies not only in the development of these models but also in their deployment, management, and scaling in production environments. With an overwhelming number of MLOps tools and packages available, practitioners often struggle to identify the most suitable solutions for their specific needs. This case study provides a comprehensive analysis of open-source MLOps tools, offering insights and comparative analysis to guide practitioners through the best options available for deploying and serving ML models efficiently. The Challenge Scale Requirements ML and AI engineers need efficient mechanisms for deploying and managing models at scale, often across multiple environments. Lifecycle Management Requirements for tracking experiments, managing model versions, and ensuring reproducibility across environments add complexity. Deployment Hurdles The deployment phase introduces challenges including scalable serving, managing dependencies, and ensuring high availability and low latency. Standardization Gaps The absence of standardized practices and tools further complicates this landscape, making the deployment and serving of ML models a difficult task. Solutions Explored To address these challenges, a variety of open-source tools have emerged, each offering unique features and capabilities to streamline the ML lifecycle. This analysis explores several key platforms in the field, examining their strengths and weaknesses for deploying and serving ML models. TensorFlow Serving • High-performance serving system designed specifically for TensorFlow models • Robust features including version management and batch processing • Out-of-the-box integration with TensorFlow ecosystem • Steep learning curve and lack of direct customer support MLflow • Versatile platform catering to end-to-end ML lifecycle • Excellent for model repository and model management • Compatible with wide range of ML libraries and deployment tools • Limited user management in self-managed instances AWS SageMaker • Fully managed service streamlining the entire ML lifecycle • Integrated Jupyter notebooks and optimized algorithms • Easy deployment capabilities for rapid model iteration • Potential for vendor lock-in and associated cost concerns Seldon Core • Powerful solution for Kubernetes environments • Supports complex inference pipelines with multiple models • Advanced features including A/B testing and model monitoring • Requires substantial Kubernetes expertise to implement BentoML BentoML bridges the gap between data science and DevOps, offering a user-friendly approach to packaging and serving ML models across frameworks. • Support for major machine learning frameworks, including MLflow and TensorFlow • High-performance API serving system for efficient model deployment • Excellent model management features for versioning and tracking • Focus primarily on model serving rather than broader ML lifecycle management Comparative Analysis Tool Best For Learning Curve Ecosystem Integration TensorFlow Serving TensorFlow-focused projects requiring high performance Steep TensorFlow-centric MLflow End-to-end ML lifecycle management Moderate Highly versatile AWS SageMaker All-in-one managed service Moderate AWS-focused Seldon Core Complex inference pipelines in Kubernetes Very Steep Kubernetes-native BentoML Streamlined model packaging and serving Gentle Framework-agnostic Key Takeaways 1 No One-Size-Fits-All The ideal MLOps tool depends heavily on your specific use case, infrastructure, and team expertise. 2 Consider Complexity Balance sophisticated features with implementation complexity when choosing a deployment solution. 3 Ecosystem Compatibility Tools that work well with your existing ML frameworks and infrastructure often provide the smoothest implementation path. 4 Scalability Planning Evaluate

tools not just for current needs but for their ability to scale with your future production requirements.);
}

Portfolio Pricing

title: "Portfolio Pricing Engine | Dean Shabi", description: "Developing a portfolio-aware pricing system for risk-optimized energy trading."
}; Project Overview I led the research and development for a modular pricing engine for a utility provider, integrating risk-aware portfolio analytics into the pricing process for energy contracts. The tool empowers analysts and commercial teams to test pricing strategies across realistic portfolio scenarios, supporting better decision-making under uncertainty. The Challenge Pricing for industrial-scale energy contracts is traditionally siloed, reactive, and manually intensive. Teams struggle to: • Quantify financial risk across an evolving portfolio • Price new contracts competitively while ensuring adequate margin • Model interdependencies between existing and future supply blocks My Approach I developed a unified simulation framework that enables scenario-based testing of risk-adjusted pricing logic. This included: Key Contributions 1 Integrated Risk Metrics Embedded VaR and ES into pricing decisions to account for portfolio tail risks. 2 Portfolio-Aware Simulation Constructed realistic scenarios to evaluate pricing strategies under uncertainty. 3 Interactive Analysis Tool Built a Streamlit-based app for users to compare pricing strategies in real time. Business Impact Key Metrics 95% Reduction in pricing time 12% Increase in portfolio margin 3x More pricing scenarios analyzed Strategic Outcomes Portfolio Intelligence Enabled nuanced pricing decisions by quantifying inter-block dependencies and forecasting risk. Operational Efficiency Streamlined decision-making through automation and scenario testing tools.);
}

Robot Failure

title: "Robot Failure Detection | Dean Shabi", description: "Machine learning models to detect and predict robot failures in automotive manufacturing."
}; Project Overview At Datamole AI, I implemented advanced anomaly detection algorithms to identify and predict robot failures in automotive manufacturing. This system monitors complex robotic systems in real-time, detecting subtle patterns that indicate potential failures before they occur. Working closely with industry specialists, our team developed custom AI solutions that analyzed multivariate sensor data from industrial robots to dramatically reduce downtime and maintenance costs. Technical Approach The robot failure detection system involved several technical components: Real-time Data Processing Pipeline to handle high-frequency multivariate signals from robot sensors in real-time. Anomaly Detection Models Advanced algorithms using both supervised and unsupervised approaches to detect deviations from normal operation. Feature Extraction Time series feature extraction techniques to identify subtle patterns preceding failures in complex sensor data. Alert System Automated alert system with configurable thresholds for different failure types and severity levels. The system employed a hybrid approach combining statistical methods, deep learning, and domain knowledge to achieve high accuracy in industrial environments with complex noise patterns. Implementation Challenges Noisy Data Working with noisy, high-dimensional sensor data from industrial environments required sophisticated filtering techniques. False Positives Balance Balancing false positives (unnecessary maintenance) with false negatives (missed failures) to optimize reliability. Model Generalization Developing models that could generalize across different robot types and configurations in varied manufacturing environments. Interpretable Results Creating interpretable results that maintenance teams could act upon without requiring data science expertise. Business Impact 1 Reduced Downtime Unplanned downtime in automotive manufacturing lines was reduced by over 35%. 2 Predictive Maintenance Identified maintenance needs before catastrophic failures occurred, preventing costly production stoppages. 3 Cost Reduction Decreased maintenance costs by enabling targeted, preventive interventions instead of major repairs. 4 Equipment Lifespan Extended robot equipment lifespan through early intervention and optimized maintenance schedules. 5 Production Quality Improved production throughput and quality by ensuring consistent robot performance. System Architecture Sensor Data Collection Signal Processing Feature Extraction Normalization Anomaly Detection Models Alert System Maintenance Interface);
}

Streamlit Guide

title: "Streamlit Guide | Dean Shabi", description: "Complete guide to developing data apps with Streamlit."
}; Project Overview Created a comprehensive guide for developing data applications with Streamlit, showcasing best practices and implementation techniques. The guide covers everything from basic setup to advanced deployment, helping data scientists and developers create professional data apps efficiently.);
}

Wind Forecasting

title: "UK's Wind Energy Generation | Dean Shabi", description: "Forecasting UK's offshore and onshore wind power generation with advanced machine learning models."
}; Project Overview This project tackles the challenge of forecasting wind power generation in the UK, encompassing both offshore and onshore wind farms to support the country's clean energy transition. Wind energy plays a critical role in this transition by offering a clean and abundant alternative to fossil fuels, with the UK emerging as a leader in wind energy adoption. UK Wind Energy Landscape The UK has emerged as a leader in wind energy adoption, boasting over 14GW of installed capacity and a staggering 23GW planned for the future. However, one key hurdle remains: accurately predicting wind power output. 14+ GW Current Capacity 23 GW Planned Capacity Wind Energy Sustainable power source with variable output The Challenge of Wind Power Prediction The very nature of wind makes it notoriously difficult to predict. Wind power generation fluctuates constantly due to a complex interplay of factors: Weather Patterns Complex and rapidly changing weather systems affect wind speed and direction across the region. Geographical Location Different coastal and inland locations experience distinct wind patterns requiring local modeling. Temporal Factors Time of day, season, and longer-term climate patterns all influence wind generation capability. This unpredictability poses a significant challenge for grid management and ensuring a stable and reliable energy supply, making accurate forecasting critical for the energy sector. Our Solution Our project addresses this challenge by developing an advanced machine learning model specifically designed for day-ahead wind power forecasting. This innovative model goes beyond existing solutions, like those used by Elexon, by incorporating a bias-correcting linear model. Technical Approach 1 Multi-variable data integration from weather and historical generation 2 Advanced time series modeling with specialized neural networks 3 Bias-correcting linear model to improve accuracy 4 Regional models for onshore and offshore wind farms Performance Improvement 57% Accuracy improvement over Elexon's forecasting system Standard Forecasting Our Solution);
}