ESG Investment Performance Analysis

Introduction

This report presents a comprehensive data analysis of ESG (Environmental, Social, Governance) investment performance. Using S&P 500 ESG Risk Ratings combined with stock data from Apple, Tesla, Microsoft, and the S&P 500 index, we investigate the relationship between ESG metrics and stock performance.

Data Preparation

- Collected ESG risk ratings data from Kaggle
- Obtained historical stock data for Apple, Tesla, Microsoft, and S&P 500
- Cleaned data by removing missing values and unnecessary columns
- Merged stock performance data with ESG metrics for analysis

Correlation Analysis: ESG Scores vs Returns

Key Findings

- **Weak Correlation with Daily Returns**: All ESG factors (Total ESG Risk, Environment, Social, Governance, Controversy scores) show correlation coefficients near zero (0.00-0.01) with daily stock returns.
- **Strong Correlations Among ESG Factors**:
  - Total ESG Risk Score strongly correlates with Social Risk Score (0.98) and ESG Risk Percentage (1.00)
  - Environment Risk Score correlates with Social Risk Score (0.94) and ESG Risk Percentage (0.86)
  - Controversy Score correlates with Social Risk Score (0.94) and ESG Risk Percentage (0.98)
- **Governance Risk Independence**: Governance Risk Score shows weak correlation with other ESG metrics and moderate negative correlation (-0.34) with Environment Risk Score.

Implications

- ESG metrics appear ineffective as signals for short-term (daily) trading strategies
- ESG factors show strong interconnections with each other, particularly between social and environmental dimensions
- Governance risk operates somewhat independently from other ESG dimensions

Return Distribution Analysis: High vs Low ESG

The histogram analysis comparing high ESG and low ESG stocks revealed:

- **Volatility Differences**: High ESG stocks (green distribution) demonstrate slightly lower volatility than low ESG stocks (red distribution)
- **Central Tendency**: Both distributions center around zero daily return, with high ESG stocks showing more frequent near-zero returns
- **Tail Behavior**: Low ESG stocks exhibit slightly fatter tails, indicating more frequent extreme returns (both positive and negative)

Regression Analysis: ESG Impact on Returns

An Ordinary Least Squares regression model yielded the following results:

- **No Explanatory Power**: R-squared and adjusted R-squared values were near zero (-0.000)
- **No Statistical Significance**: F-statistic (0.1109) with p-value of 1.00 indicates the model lacks significance
- **No Significant ESG Variables**: All ESG variables showed high p-values (>0.05)
- **Severe Multicollinearity**: Condition number of 9.53e+14 indicates highly correlated independent variables

Multicollinearity Analysis (VIF)

Variance Inflation Factor analysis showed:

- Total ESG Risk Score: VIF = 20,220 (severe multicollinearity)
- Environmental Risk Score: VIF = 6,325 (very high collinearity)
- Governance Risk Score: VIF = 2,717 (high collinearity)
- Social Risk Score: VIF = 82.3 (high collinearity)
- Controversy Score: VIF = 0.00 (no collinearity issue)

Feature Engineering & Advanced Modeling

To improve predictive performance, we explored:

1. **Principal Component Analysis (PCA)**: Reduced dimensionality of ESG variables
2. **Random Forest Model**: Initial R^2 score: -0.001453
3. **Neural Network Model**: Initial R^2 score: -0.001108
4. **Feature Engineering Improvements**: After refining features and model parameters, achieved best R^2 score of 0.6158 (61.58% accuracy)

Feature Importance Analysis

Analysis of Random Forest feature importance revealed:

- **Technical Indicators Dominate**:
  - Momentum (highest importance)
  - Volatility (second highest)
  - EMA_200, RSI, MACD_Signal, MACD, Bollinger Bands
- **Limited Impact of Fundamental & ESG Factors**:
  - EPS, Book-to-Market, ROE, P/E Ratio, Debt-to-Equity, Current Ratio, ROA, Market Cap
  - Total ESG Risk score showed minimal importance in the model

Model Performance Comparison

Model	R^2 Score	Mean Squared Error (MSE)
Linear Regression	~0.68	~0.00135
Neural Network	~0.66	~0.0014
Random Forest	~0.62	~0.0016
Gradient Boosting	~0.53	~0.00195

Prediction Error Analysis

- **Neural Network**: More concentrated error distribution around zero
- **Random Forest**: Wider error distribution with heavier tails
- Both models centered around zero error, indicating no systematic bias

Time Series Prediction Performance

Time series analysis of predicted vs. actual returns showed:

- All models generally tracked the direction of actual returns
- Models produced smoother predictions than actual returns
- All models struggled to capture extreme market movements
- Gradient Boosting showed the smoothest predictions with largest deviations
- Neural Network and Random Forest provided more responsive predictions

Trading Strategy Backtest Results

Four strategies were backtested:

1. **Buy & Hold**: Showed steady growth with dramatic increase near end of testing period
2. **Mean Reversion**: Limited growth, staying relatively flat
3. **Momentum**: Limited initial growth with noticeable upward trend at end of period
4. **Threshold**: Worst performer with minimal returns or slight losses

Monte Carlo Simulation

Monte Carlo simulation of portfolio value over 250 trading days showed:

- Wide range of potential outcomes ($50,000 to $400,000 from $100,000 starting value)
- General upward trend in many simulations
- Increasing dispersion over time, highlighting compounding uncertainty
- Concentration around initial value in early stages

Conclusions

1. **ESG and Daily Returns**: ESG metrics show minimal correlation with daily stock returns, suggesting limited utility for short-term trading strategies.

2. **Risk Profile Differences**: High ESG stocks demonstrate slightly lower volatility than low ESG stocks, potentially offering more stability.

3. **Model Performance**: Advanced models (Neural Network, Random Forest) can achieve reasonable predictive performance (R^2 ~0.62-0.66) when combining ESG metrics with technical indicators.

4. **Feature Importance**: Technical indicators dominate model predictions, with ESG metrics showing minimal direct influence on short-term returns.

5. **Trading Strategy Effectiveness**: Buy & Hold outperformed other strategies in the backtest, with Momentum showing promise in trending markets.

Recommendations

1. **Long-term Focus**: Consider ESG factors for long-term investment horizons rather than short-term trading signals.

2. **Risk Management**: Use ESG scores as potential indicators of volatility and stability rather than return predictors.

3. **Model Improvement**: Continue refining predictive models by:
    - Exploring non-linear relationships
    - Using longer time horizons for returns (weekly/monthly)
    - Combining technical and ESG factors strategically
4. **Feature Engineering**: Focus on creating more sophisticated composite features from ESG data that might better capture their relationship with financial performance.
5. **Strategy Development**: Consider ESG metrics primarily for risk management and portfolio diversification rather than alpha generation in short-term strategies.