**Forecasting EBITDA/EV for Enhanced Portfolio Construction: A Multi-Industry Machine Learning Approach**

**1. Introduction**

This research investigates the application of machine learning techniques to forecast company fundamentals across different industries for the purpose of portfolio construction. Traditional quantitative investment strategies typically rely on current financial metrics to identify undervalued securities. However, the long-term performance of an investment is ultimately determined by a company's future financial performance rather than its current state.

Building upon the groundbreaking work of Alberg and Lipton (2017), this study extends their approach by examining how the predictive power of machine learning models varies across different industries. While Alberg and Lipton demonstrated that forecasting fundamentals using neural networks could enhance overall portfolio returns, our research examines whether this effect is consistent across all sectors or if certain industries exhibit greater predictability and, consequently, higher potential returns.

The research utilizes deep learning architectures, specifically Recurrent Neural Networks (RNNs) and Multilayer Perceptrons (not too sure yet, will need read more paper and consult with experts), to forecast future EBITDA (Earnings Before Interest, Taxes, Depreciation, and Amortization) values based on historical financial data. These forecasted values are then used to calculate the EBITDA/EV (Enterprise Value) ratio, a relative value factor for portfolio construction. By comparing the performance of industry-specific models against general models, we aim to identify which sectors benefit most from this forecasting approach and which model serve as the strongest predictors for each industry.

**2. EBITDA/EV as a Value Factor for Portfolio Construction**

**2.1 Definition and Significance**

The EBITDA/EV ratio is a valuation metric that compares a company's earnings before interest, taxes, depreciation, and amortization to its enterprise value. Enterprise value represents the total value of a company, calculated as market capitalization plus debt minus cash and cash equivalents.

This ratio is particularly valuable for several reasons:

1. **Operating Focus**: By using EBITDA, the ratio focuses on core operating performance before the effects of capital structure, tax environments, and non-cash expenses.
2. **Capital Structure Neutrality**: Unlike price-to-earnings (P/E) ratios, EBITDA/EV accounts for differences in capital structure, making it more suitable for comparing companies with varying debt levels.
3. **Industry Comparability**: The ratio facilitates comparisons across different industries, as it excludes the effects of financing and accounting decisions that might vary by sector.
4. **Cash Flow Proxy**: EBITDA serves as a reasonable proxy for operating cash flow, providing insight into a company's ability to generate cash from operations.

**2.2 Empirical Evidence**

(I’ve affiliated a paper that argues EBITDA/EV is the best value factor, but I think probably more papers would be needed to argue why it’s the best despite its portfolio’s performance, the above are just some points I listed, and more should be included in both 2.1, 2.2 in the final paper)Extensive academic research has demonstrated the effectiveness of the EBITDA/EV ratio in identifying undervalued securities. Studies have shown that portfolios constructed by ranking stocks according to this ratio have historically outperformed the market and other traditional value metrics. The ratio effectively identifies companies generating strong operating earnings relative to their total value, which often indicates undervaluation.

According to Alberg and Lipton's research, a strategy based on the EBITDA/EV factor achieved a compound annualized return (CAR) of 14.4% from 2000-2016, significantly outperforming both the S&P 500 (4.5%) and the market average (7.7%). This performance reinforces the ratio's effectiveness as a stock selection criterion.

**3. The Importance of Forecasting EBITDA/EV**

**3.1 Theoretical Foundation**

While traditional value investing approaches use current financial metrics to identify potentially undervalued securities, the actual investment returns will ultimately depend on future company performance. Alberg and Lipton demonstrated through simulation that a clairvoyant model with access to future fundamentals could achieve exceptional returns—as high as 44% annual returns when using 12-month future EBITDA/EV values.

This finding establishes a clear theoretical foundation: if we can accurately forecast future fundamentals, particularly EBITDA, we should be able to construct portfolios that outperform those based solely on current financial data.

**3.2 Practical Implementation**

In practical terms, forecasting EBITDA/EV offers several advantages over traditional approaches:

1. **Forward-Looking Perspective**: By incorporating predicted future fundamentals, the approach aligns more closely with the theoretical basis of investing—that present value reflects expected future cash flows.
2. **Capturing Trend Information**: The forecasting models can identify companies whose operating performance is improving but not yet reflected in market prices.
3. **Reducing Value Traps**: Companies with currently strong metrics but deteriorating fundamentals (value traps) can be identified and avoided through forward-looking analysis.
4. **Exploiting Market Inefficiencies**: Markets may not fully incorporate predictable changes in fundamentals, creating opportunities for excess returns.

The empirical evidence supports this approach. Alberg and Lipton's lookahead factor models achieved a CAR of 17.1% (using MLP) and 16.7% (using LSTM) compared to 14.4% for traditional factor models, representing an improvement of nearly 3 percentage points annually.

**3.3 Industry-Specific Considerations**

Our research extends this approach by examining whether the benefits of forecasting EBITDA/EV vary systematically across industries. There are several reasons to expect industry-level differences (more paper should be read and referenced):

1. **Earnings Predictability**: Some industries have more stable and predictable earnings patterns than others.
2. **Industry-Specific Business Cycles**: Different sectors may experience varying sensitivity to economic cycles, affecting the predictability of their fundamentals.
3. **Competitive Dynamics**: Industries with high barriers to entry may maintain more stable earnings compared to highly competitive sectors.
4. **Regulatory Environment**: Heavily regulated industries might demonstrate different patterns of fundamental predictability.

By training industry-specific models, we aim to identify which sectors benefit most from this forecasting approach and tailor the investment strategy accordingly.

**4. Data and Methodology**

**4.1 Data Sources**

This research utilizes data from the following sources(things the I am currently working on, this will take considerable amount of time):

1. **CRSP (Center for Research in Security Prices)**: Monthly stock return data, prices, and trading volume for all stocks listed on NYSE, NASDAQ, and AMEX.
2. **Compustat North America**: Quarterly financial statement data including income statements, balance sheets, and cash flow statements.
3. **CRSP-Compustat Merged Database**: Linking table that connects CRSP and Compustat identifiers.

The study includes all publicly traded US companies from January 2010 to December 2015(will probably be adjusted later), excluding financial firms (due to their unique accounting practices) and companies with inflation-adjusted market capitalizations below $100 million(more constraints could be added later, these are from Alberg and Lipton's paper).

**4.2 Feature Engineering**

Following Alberg and Lipton's methodology but extending it with industry-specific considerations, we engineer a comprehensive set of features from financial statements. For each company at each time point, we use a trailing window of 5 years of quarterly data to predict the EBITDA 12 months into the future.

**5. Financial Features and Their Significance**

The following financial metrics have been selected for feature engineering, with explanations of their relevance to EBITDA forecasting and portfolio construction(reference: <https://wrds-www.wharton.upenn.edu/data-dictionary/comp_na_daily_all/fundq/> ):

**5.1 Income Statement Metrics**

1. **Revenue (revtq)**: The total amount of money generated by a company's business activities. Revenue is the fundamental starting point for EBITDA calculation and captures a company's ability to attract customers and set prices. Revenue growth trends are strong indicators of future EBITDA potential.
2. **Cost of Goods Sold (cogsq)**: Direct costs attributable to the production of goods or services sold. COGS trends reflect production efficiency and input cost management, directly affecting gross margins and consequently EBITDA.
3. **Selling, General & Administrative Expense (xsgaq)**: Overhead costs not directly tied to production. SG&A management is crucial for operational efficiency and represents a significant portion of controllable costs that impact EBITDA.
4. **Operating Income (oiadpq)**: Profit from core business operations before interest and taxes. This closely approximates EBITDA before depreciation and amortization and thus provides direct insight into its trajectory.
5. **Operating Income Before Depreciation (oibdpq)**: EBITDA minus non-operating income. This metric closely tracks EBITDA and provides valuable information for forecasting future values.
6. **Depreciation (dpq)**: Non-cash expense allocating the cost of tangible assets over their useful lives. Understanding depreciation patterns helps forecast the difference between operating income and EBITDA.
7. **Operating Cash Flow (oancfy)**: Cash generated from normal business operations. While EBITDA is an approximation of operating cash flow, examining the relationship between the two provides insight into earnings quality.
8. **Net Income (niq)**: Total earnings after all expenses. Net income trends can signal overall business health and sustainability of operations that generate EBITDA.

**5.2 Balance Sheet Metrics**

1. **Total Assets (atq)**: Everything a company owns with economic value. Asset base size and growth inform operational scale and future earnings potential.
2. **Total Liabilities (ltq)**: All debts and obligations. Liability levels affect a company's financial flexibility and enterprise value calculations.
3. **Common Equity (ceqq)**: Shareholders' residual claim on assets. Equity levels reflect historical profitability and investment in the business.
4. **Cash and Equivalents (cheq)**: Highly liquid assets. Cash positions affect enterprise value calculations and indicate financial flexibility.
5. **Accounts Receivable (rectq)**: Money owed by customers. Receivables management affects cash conversion efficiency and may signal revenue recognition issues.
6. **Inventories (invtq)**: Goods held for sale or production. Inventory management reflects operational efficiency and working capital utilization.
7. **Property, Plant & Equipment (ppentq)**: Tangible long-term assets. Capital intensity affects depreciation and capital expenditure requirements, impacting EBITDA.
8. **Current Assets/Liabilities (actq/lctq)**: Short-term resources and obligations. Working capital management affects operational efficiency and cash flow.
9. **Long-term Debt (dlttq)**: Borrowings due beyond one year. Debt levels impact enterprise value calculations and interest expenses.
10. **Common Shares Outstanding (cshoq)**: Number of shares issued minus treasury stock. Used in market capitalization calculations for enterprise value.

**5.3 Derived Financial Ratios**

1. **Gross Margin**: (Revenue - COGS) / Revenue. Measures pricing power and production efficiency, fundamental drivers of EBITDA.
2. **Operating Margin**: Operating Income / Revenue. Directly indicates a company's operational efficiency and profitability.
3. **Asset Turnover**: Revenue / Total Assets. Measures how efficiently a company uses its assets to generate revenue, affecting EBITDA generation.
4. **Inventory Turnover**: COGS / Inventories. Reflects inventory management efficiency, particularly important in retail and manufacturing.
5. **Current Ratio**: Current Assets / Current Liabilities. Indicates short-term financial health and liquidity.
6. **Debt-to-Equity**: Total Liabilities / Stockholders' Equity. Reflects financial leverage, affecting both risk and enterprise value.
7. **Interest Coverage**: Operating Income / Interest Expense. Measures a company's ability to meet interest obligations, indicating financial stability.

There should be more solid criteria to select features, so this sections would be updated