**Deep Robust RL for Practical Algorithmic Trading – only handle one asset per agent**

Data Extraction

* Three effective methods to filter financial time series and eliminate most of the uncertainty noise
* Use SDAE to make the model more robust

**Method 1** – Sampling random length of the episode to increase the model’s generalization and exploration

**Method 2**- Reducing impacts of news and notices. We reduce the impact of these news and notices by specific setting. For example, as most news and notices of quoted companies are released off the trading time in China, their impacts usually occur in the opening time (high open or low open). According to this phenomenon, we extract the financial time series within the trading period (9:30 am to 11:30 am and 13:00 pm to 15:00 pm)

**Method 3** - The low volatility of the financial time series will have a detrimental effect on our predictions due to their abnormal fluctuations. The low volatility of time series is mainly caused by the individual investors (not institutional investors) which can be regarded as noise. The market is inactive accompanied by a lot of noise at those time points, thus we remove those series with low volatility in order to reduce noise and unsteadiness.

\*Introduce private variables to increase difference between states – remaining cash + sharp ratio

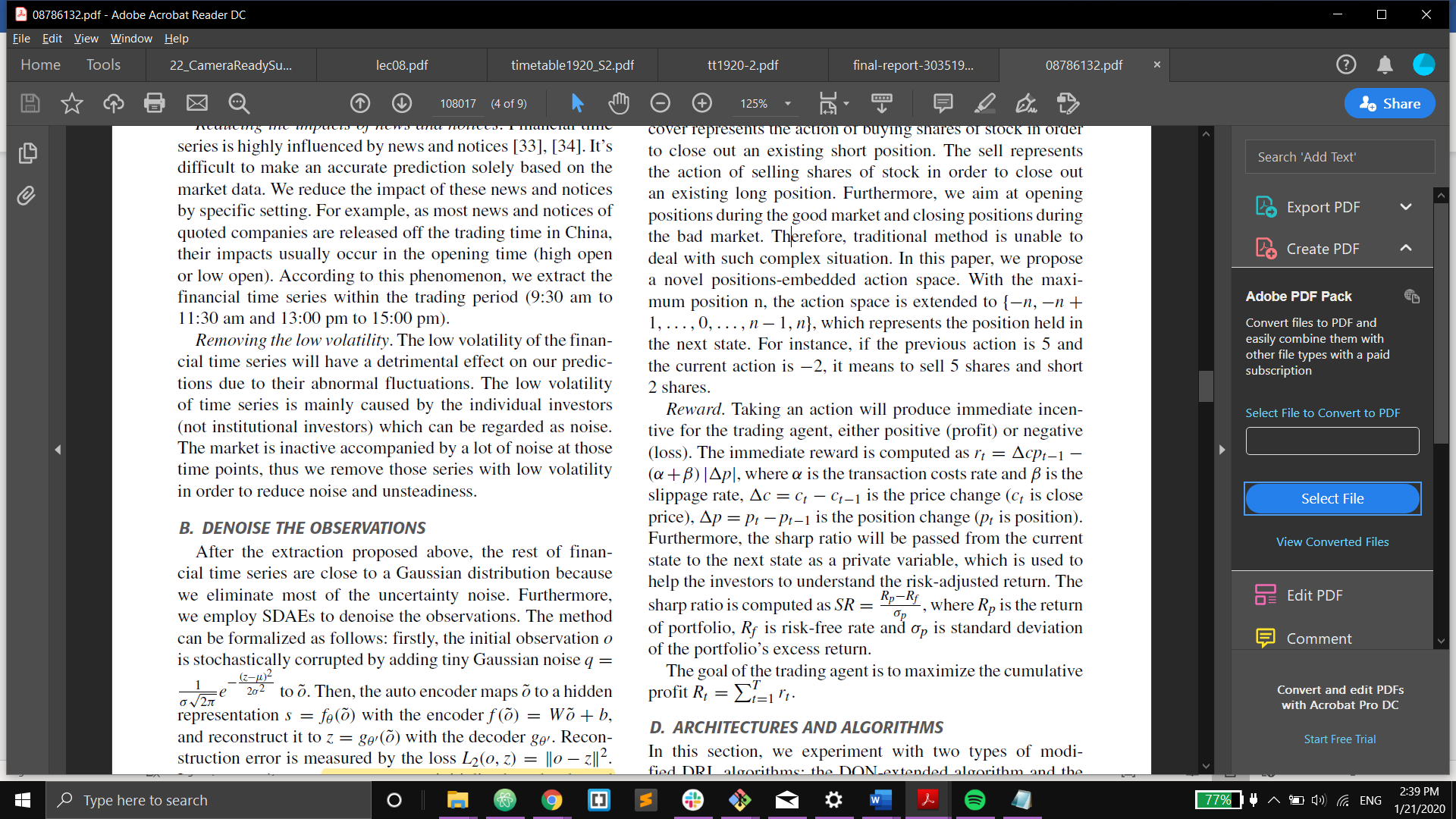
\*Data is passed to SDAE after applying the three methods

RL Formulation for trading

1. State representation (All normalized)
   1. Market variables: open, close, high, low price and trading volume
   2. Technical Indicators: MACD, MA, EMA, ATR, ROC (See W. Bao, J. Yue, and Y. Rao, “A deep learning framework for financial time series using stacked autoencoders and long short-term memory'”)
   3. Private Variables: remaining trading cash, previous sharp ratio, which represent how much cash has been left and how much prot or loss has been got
2. Action space

the real trading environment is more complex, where exists a great many operations corresponding to different trading directions (long, sell, short, cover). The long and the short is equal to buy and sell, respectively. The cover represents the action of buying shares of stock in order to close out an existing short position. The sell represents the action of selling shares of stock in order to close out an existing long position. Furthermore, we aim at opening positions during the good market and closing positions during the bad market. Therefore, traditional method is unable to deal with such complex situation. In this paper, we propose a novel positions-embedded action space. With the maximum position n, the action space is extended to {-n, -n+1, …, 0, …, n-1, n} , which represents the position held in the next state. For instance, if the previous action is 5 and the current action is -2, it means to sell 5 shares and short 2 shares.

1. Reward

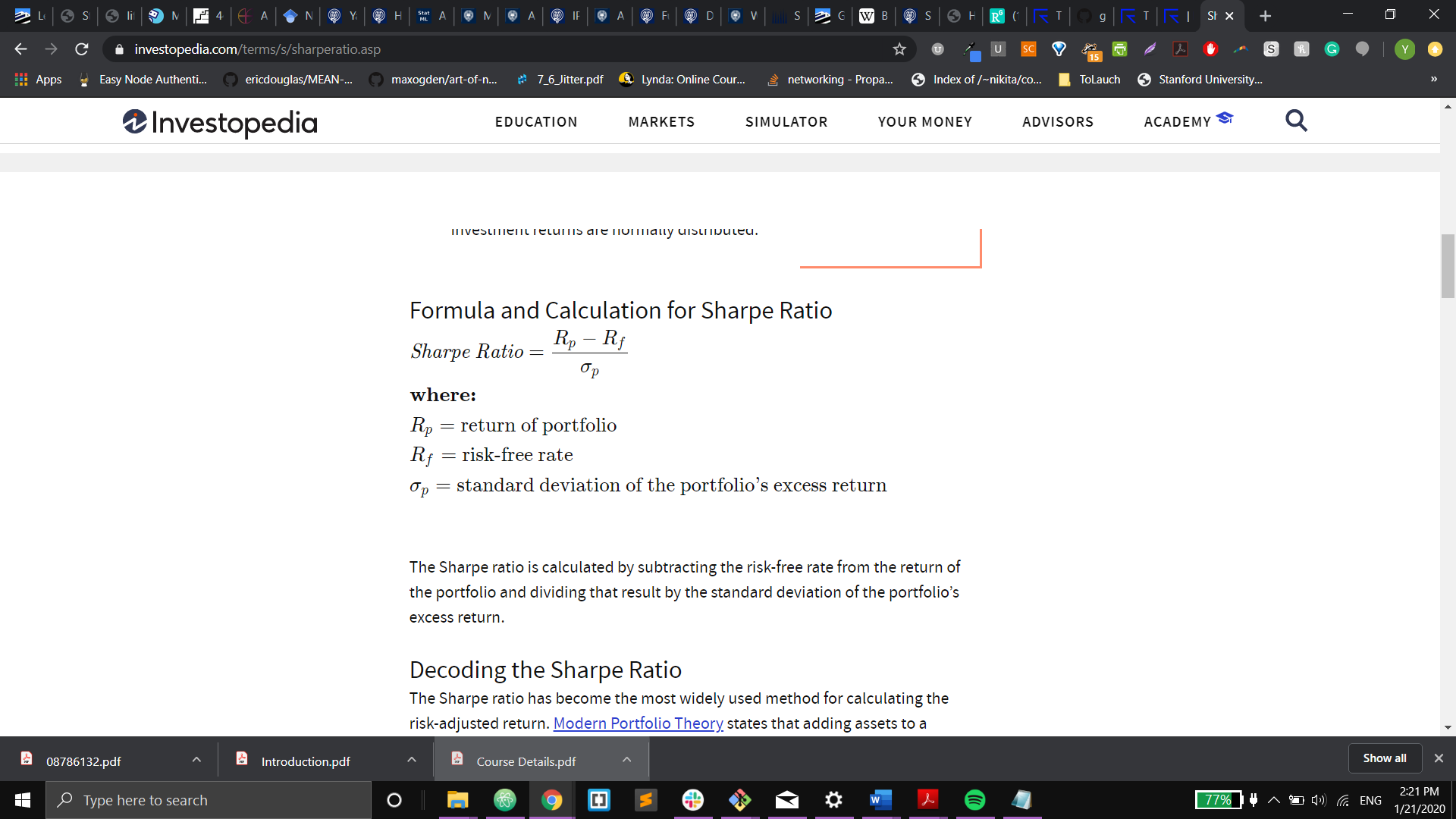


1. Evaluation Metric

All models and strategies are evaluated by the metric annualized return (AR) and the metric sharp ratio (SR). The annualized return is the geometric average of the money earned by an investment each year over a given time period, and SR is computed as mentioned above.

Glossary

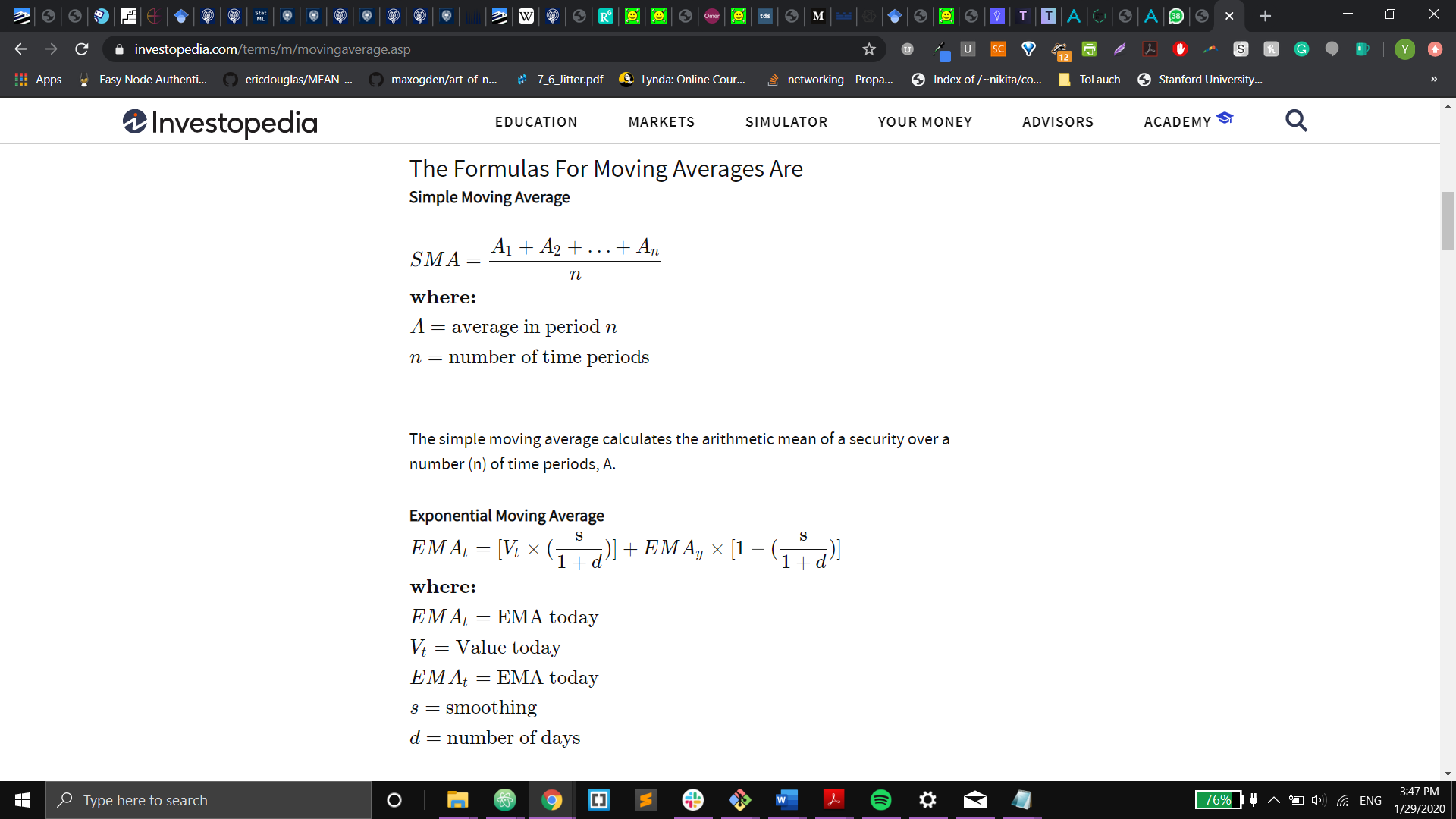
1. Sharpe Ratio



1. Moving Average Convergence Divergence (MACD)

Moving Average Convergence Divergence (MACD) is a trend-following momentum indicator that shows the relationship between two moving averages of a security’s price. The MACD is calculated by subtracting the 26-period Exponential Moving Average (EMA) from the 12-period EMA. The result of that calculation is the MACD line. A nine-day EMA of the MACD called the "signal line," is then plotted on top of the MACD line, which can function as a trigger for buy and sell signals. So we have tom compile 26 and 12-period EMA first

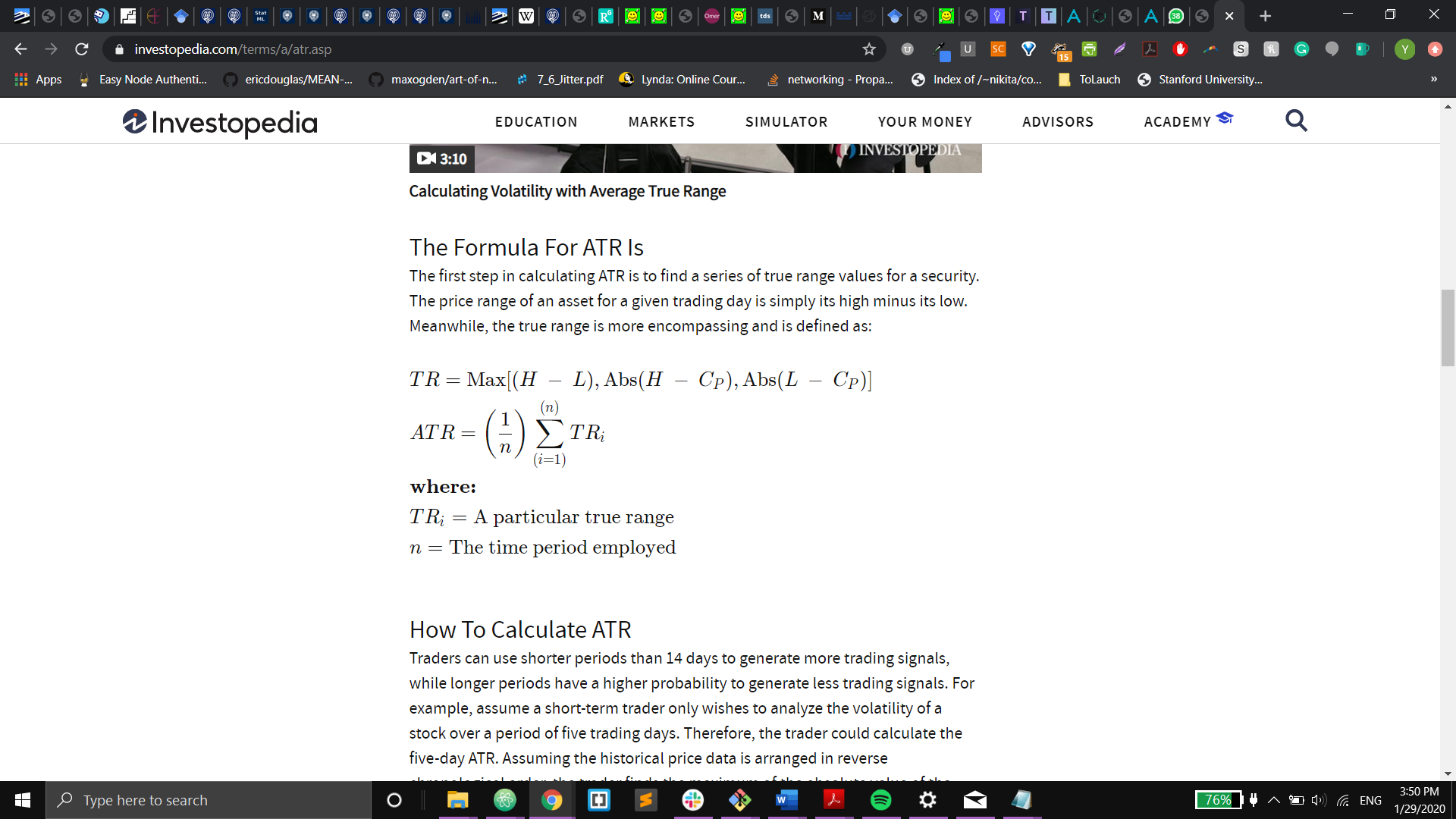
1. Moving Averages



1. Average True Range (ATR) – measures market volatility

The average true range (ATR) is a technical analysis indicator that measures market volatility by decomposing the entire range of an asset price for that period. Specifically, ATR is a measure of volatility introduced by market technician J. Welles Wilder Jr. in his book, "New Concepts in Technical Trading Systems."

The true range indicator is taken as the greatest of the following: current high less the current low; the absolute value of the current high less the previous close; and the absolute value of the current low less the previous close. The average true range is then a moving average, generally using 14 days, of the true ranges.

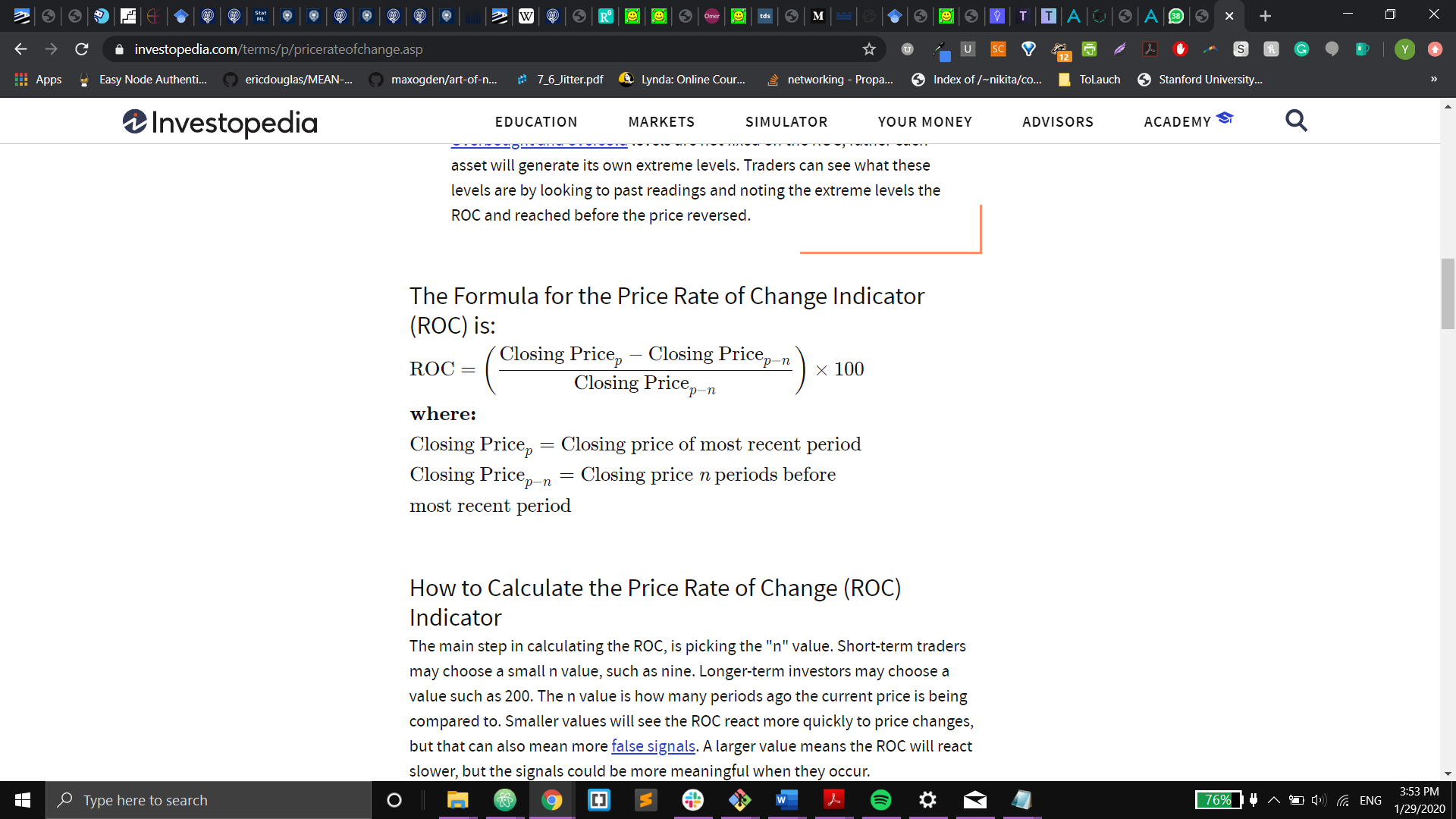


Traders can use shorter periods than 14 days to generate more trading signals, while longer periods have a higher probability to generate less trading signals. For example, assume a short-term trader only wishes to analyze the volatility of a stock over a period of five trading days. Therefore, the trader could calculate the five-day ATR. Assuming the historical price data is arranged in reverse chronological order, the trader finds the maximum of the absolute value of the current high minus the current low, absolute value of the current high minus the previous close and the absolute value of the current low minus the previous close. These calculations of the true range are done for the five most recent trading days and are then averaged to calculate the first value of the five-day A

1. Price Rate of Change (ROC)

The Price Rate of Change (ROC) is a momentum-based technical indicator that measures the percentage change in price between the current price and the price a certain number of periods ago. The ROC indicator is plotted against zero, with the indicator moving upwards into positive territory if price changes are to the upside, and moving into negative territory if price changes are to the downside.

The indicator can be used to spot divergences, overbought and oversold conditions, and centerline crossovers.



The main step in calculating the ROC, is picking the "n" value. Short-term traders may choose a small n value, such as nine. Longer-term investors may choose a value such as 200. The n value is how many periods ago the current price is being compared to. Smaller values will see the ROC react more quickly to price changes, but that can also mean more false signals. A larger value means the ROC will react slower, but the signals could be more meaningful when they occur.

Select an n value. It can be anything such as 12, 25, or 200. Short-term trader traders typically use a smaller number while longer-term investors use a larger number.

Find the most recent period's closing price.

Find the period's close price from n periods ago.

Plug the prices from steps two and three into the ROC formula.

As each period ends, calculate the new ROC value.