Wisdom of the Crowd: Ensemble Methods for Determining Market States

Abstract— **This paper aims at discerning the hidden state of the market: bull, bear, or static. Three models (Hidden Markov, Markov Switching, and Neural Network) and two strategies (simple long- or short-only and a vol-timing strategy) are explored, with model output being combined using the ensemble majority voting methodology. The ensemble method on the vol-timing strategy produces the highest Sharpe ratio. This underpins the notion that an ensemble technique can yield better performance than the constituent models for the purposes of predicting the market state.**

***Index Terms*— Hidden Markov Model, Neural Network, Markov Regime Switching Model, Ensemble Methods**

I. INTRODUCTION

The tendency of financial markets to change its behaviour over time, often cyclically, has led many academics and practitioners to posit the existence of different stock market *regimes*. Usually caused by changes in expectation of future cash flows and interest rates, these market regimes are believed to precede the business cycle and fortell of economic booms and busts. *Bull* markets are characterized by rising prices and mild volatility, and *bear* markets are characterized by falling prices and elevated volatility, with *static* markets falling in-between these two.

As it relates to decision-making and economic analysis, the state of the equity market is an important variable. The utility of discerning past, current, and future regimes is immense and can have several applications: informing investor strategies and risks (as in [7] whereby smaller, lesser collateralized firms are more strongly affected by a recession state than larger ones), correlation with risk premia (see [3]), and forecasting of macro-economic variables (to name a few).

Utilizing market regimes to inform investment decisions is challenging because the state of the equity market is *latent*: unobserved and hidden. As such, one needs to utilize specialized models to infer and identify the hidden market state. Moreover, evaluating model performance and establishing criteria for gradating models remains a concern. Nevertheless, the endeavour remains fruitful.

In this paper we explore several different methods for inferring the hidden regime state (bull, bear, or static) of the market. We employ and study Hidden Markov Models, Markov Regime Switching Models, and Neural Network Models (using labeled data created by Rule-Based methods) as means to ascertain the equity market state and then use an ensemble method to systematically aggregate the information contained in these approaches and improve results.

We evaluate model performance via proxy variables: the profit, Sharpe ratio, maximum drawdown and skewness of a hidden-state trading strategy compared to a buy and hold strategy for the years 2018 to 2021. The majority voting ensemble method was optimal in almost all cases. Thus, with models as with crowds, collective knowledge is superior to the knowledge of the few.

A. LITERATURE REVIEW

Three main methods of regime identification have been explored in the literature in the last two decades. First, changes in risk aversion of market players could signify changing market regime dynamics. In [5], risk aversion indicators are found to be good leading indicators for stock market crises (and, thus, bear markets). More generally, the idea is that latent market regime variables could be inferred via observable values which are proxies for risk aversion.

Second, the characteristics of time series data (mean, variance, etc.) may vary systematically across market states. Markov Regime Switching models propose that the observed time-series data falls into different, recurring regimes. Several papers, [2] and [8], explore these models, inferring the probabilities of latent (market) state variables.

Lastly, rule-based methods can be used to procedurally label bull and bear markets. Rules can be based on statistical properties of market returns or other key economic data. These methods have been shown to be preferable for in-sample identification, with regime-switching models performing better for out-of-sample forecasting (see [15]).

B. OUTLINE

Section II introduces and briefly explains the models and methods we considered in our study. Section III describes the data and its sources. Section IV outlines our approach and our hidden-state trading strategy, i.e. the strategy employed which has the predicted market regime as input. Lastly, in Section V & VI, we display our results and draw conclusions.

II. MODELS & METHODS

In this section, we briefly introduce the models studied for predicting the market's hidden state and the methods used to augment our approach. The specific model construction and data inputs are outlined in Section IV.

A. HIDDEN MARKOV MODEL

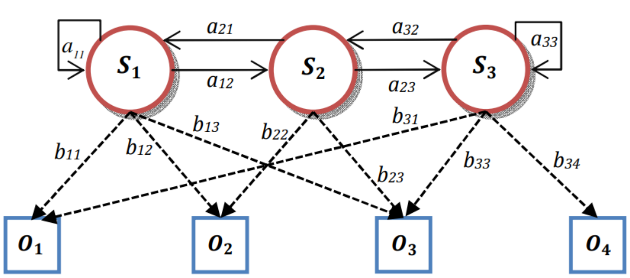

Fig. 1. Three-State Hidden Markov Model. Hidden States = $\{S_1, S_2, S_3\}$, transition matrix A, emission matrix B, Observations = $\{O_1, O_2, O_3, O_4\}$. Image Source: [6]

With reference to Fig. 1, Hidden Markov Models (HMM) are models wherein the system (S_n, O_n) is assumed to be a Markov process (i.e. probability of each event depends only on the previous event) with unobservable states S_n and an observable process O_n. Our goal is to learn about S_n by observing O_n. The transition matrix A represents the probability of moving between hidden states, and emission matrix B is the probability of observing O_i given we are in state j.

With reference to our market regime problem, we have three hidden states $S = \{bull, bear, static\}$ and full control on which observable process(es) to consider. Given the observable processes, we first want to estimate the transition and emission probabilities that are most likely to give O using the training data. That is to find,

$$A^*, B^* = \arg\max_{A,B} P(O|A,B)$$

Then, given A, B and a given sequence O, we need to find the hidden sequence S that is most likely to generate O, that is to find

$$S^* = \arg\max_{S} P(O|S,A,B)$$

More details of these algorithms can be found in [4].

B. MARKOV REGIME SWITCHING MODEL

Markov Regime Switching Models are descended from HMMs, where the objective is to estimate an exogenous variable, and not the regime itself as formulated in [9]. Parameters take the form of regime-specific values, where the regime-switching mechanism is governed by an unobservable state variable and is modeled using a (hidden) Markov model. The Markovian property dictates that the current state value (and thus, parameter values) depends only on the immediately previous state value. The model equation is,

$$y_t = \beta_{0,s_{t-1}} + \beta_{1,s_{t-1}} y_{t-1} + e_t$$

where, the coefficients of the regression are state s_t dependent. [10] extends it further by allowing for the error variance to change with regimes, thus allowing for a heteroskedasticity.

$$e_t = \mathcal{N}\left(0, \sigma^2_{s_{t-1}}\right)$$

We use the version extended by [11] which allows us to have time-varying transition probabilities between regimes. Instead of having a constant value the the transition probabilities are now modeled as,

$$p_{ij,t} = \frac{e^{x^T_{t-1}\beta}}{1 + e^{x^T_{t-1}\beta}}$$

where $p_{ij,t}$ is the probability of transitioning from regime i to regime j at time period t. This also gives us the flexibility to choose economically significant exogenous variables to model the change in probabilities.

For this model, again, we have three hidden states $S = \{bull, bear, static\}$.

C. RECURRENT NEURAL NETWORK

The problem of market regime prediction can be reformulated under the supervised-learning paradigm. If one could label the training data (through a rule-based algorithm, for example) as *bull*, *bear* or *static*, a Recurrent Neural Network (RNN; as shown in Fig. 2) could be used to learn how to characterize these states in terms of market and macro data and classify the market regimes. Note: our rule-based methodology for labelling the data will been detailed in the section IV.

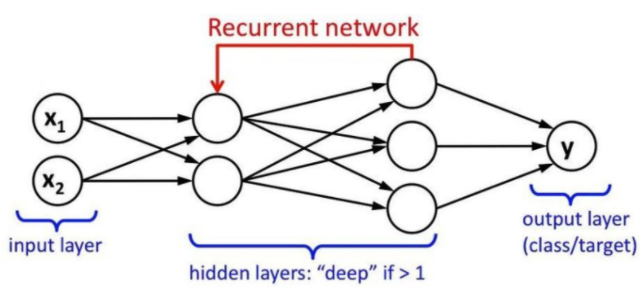

Fig. 2. Recurrent Neural Network. Image Source: [17]

RNNs are a modification of the classic feed-forward (unidirectional) neural network, wherein some function $f^*(x)$ is approximated by the function $f(x;\theta)$ by learning the parameters θ that result in the best approximation. RNNs extend these networks to include feedback connections, allowing for sequential and temporal inputs and dynamic behaviour.

However, one caveat of the basic RNN architecture is that, in practice, they can be difficult to train to solve problems involving long-term temporal dependencies (due to exploding and vanishing gradient problem). Long Short-Term Memory (LSTM) [13] networks overcome this gradient problem by introducing two gate units which ensure constant back-flow of error (Fig 3). The detail working of the gates can be found at [20]. In short, LSTMs are sophisticated networks that are well-suited to handle time series data and learn long-term temporal dependencies.

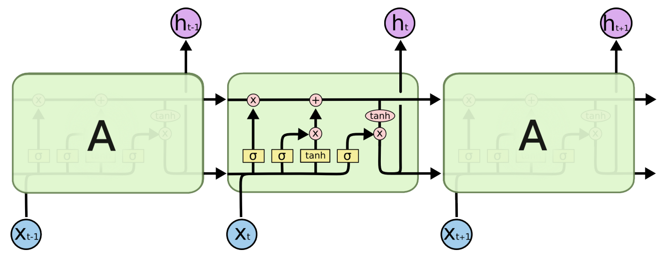

Fig. 3. Long Short-Term Memory. Image Source: [20]

D. PRINCIPAL COMPONENT ANALYSIS

Principal component analysis (PCA) is the process wherein an orthonormal basis of the coordinate space spanned by the data is constructed for the purposes of dimensionality reduction with the least information loss. High-dimensional data is often undesirable due to sparsity and increased computational complexity. As such, PCA transforms data into a low-dimensional representation while maintaining the meaningful proprieties of the original data.

We use PCA on a large set of economic and market data inputs (described in Table I) to help our models learn by reducing the noise in the data and thus reducing the opportunity for the model to overfit to the training data.

We followed the following steps to apply the PCA on our dataset:

1) Standardize the input variables and calculate the covariance matrix of the data
2) Compute the eigenvalues and eigenvectors of the covariance matrix

3) Choose top k principal components which explain at least 90% variance
4) Transform original data using the eigenvectors of selected principal components

E. ENSEMBLE LEARNING

Ensemble methods seek to combine many learning algorithms to obtain superior predictive power over any single constituent. With the finalized hidden-state models, we'll look to combine their state prediction and achieve better results. In our case, majority voting method to achieve this. We take the mode of the predictions at each timestep, which thus yields our final prediction. We map our regimes $\{bull, bear, static\}$ into numerical categories $\{1, 0, -1\}$ respectively. Our ensembled state is then,

$$S_{ens,t} = \text{Mode}\{S_{i,t}\}$$

for each model i at time t. As we have 4 models (HMM - Baseline, HMM - PCA, MSM, LSTM), there can be cases of conflict in voting. In those cases we use static state as the tie-breaker.

III. DATA

For our analysis, we collect open, high, low, close data for the Russell 3000 from Yahoo Finance and volume data using the Russell 3000 ETF ticker. From this data, we compute returns on various horizons (weekly, monthly and quarterly), open-to-close returns, and price deviation in a two week period.

Table I provides a summary of all data item and sources we used in our analysis. Most of the time series obtained were not stationary and had trends. We analyzed each of the variables using augmented Dickey-Fuller (ADF) test and then used various techniques such as log transformation and first order difference to remove the unit root from the variables.

We sample the data weekly, in order to reduce the noise from daily fluctuations. Moreover, some of the data is only available at lower frequencies in which case we forward-filled to ensure there is no missing data. For example, since GDP data is available monthly, all the weeks with no data are filled with the preceding month's GDP.

TABLE I
SUMMARY OF DATA ITEM, FREQUENCY & SOURCE

Data Item	Frequency	Source
Stock Market Indices (RUA, IVW, SP500)	Daily	FRED
Bill Yields (1, 3 & 6M)	Daily	FRED
Bond Yields (1, 2, 3, 5, 7, 10Y)	Daily	FRED
Fama French Factors (MKT, SMB, HML, RMW, CMA)	Daily	Kenneth R. French
Option-Adjusted Spread (AAA, BBB, High-Yield, Emerging)	Daily	FRED
Aggregate Indices	Daily	Squeeze Metrics
VIX & SKEW	Daily	FRED
Implied Correlation	Daily	CBOE
Sector Valuation	Monthly	WRDS
Book/Price Ratios for GICS Sectors	Monthly	Robert Shiller
Long-Term Growth Forecast	Monthly	CBO
Inflation Expectations	Monthly	Cleveland Fed
Avg. Investor Stock Holding	Quarterly	FRED
Macro Forecast Surprises (GDP, Inflation, 3M Rate)	Quarterly	Philadelphia Fed

We'll describe data segments and briefly justify their inclusion:

- **Returns** from the stock market and Fama-French factor portfolios could directly distinguish bull and bear markets from time-series properties of returns.
- **Yield** levels and changes could be associated with different market states as they reflect fiscal and monetary policy.
- **Sentiment** proxies (VIX, SKEW, Implied Correlation) measure market uncertainty and perceived tail-risk which are likely to change across market regimes.
- **Positioning** measures (i.e. GEX: gamma position of dealers) indicate whether hedging

activity is likely to reinforce or suppress market moves.
- **Valuations** at the market and sector levels can indicate market states, with high valuation multiples relative to history indicative of a bull market and low multiples associated with bear markets.
- **Risk Premium** calculated on an ex-ante basis is similar to valuation in that it measures the market's required compensation for equity risk, here taking into account real bond yields and expected long-term real profit growth.
- **Surprises** bull markets might be characterized by a string of positive economic surprises - higher-than-expected growth, lower inflation and lower rates. Bear markets would may have the opposite profile.
- **Macro data** bull markets may involve a mix of strong readings of macro variables, with bear markets seeing below-average realizations of those indicators.

IV. OUR APPROACH

In this section, we outline our model pipeline for predicting the hidden market state and the two trading strategies we employ that use these states to invest in the Russell 3000 Index.

A. BASELINES

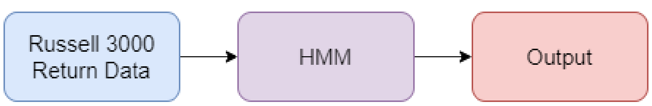

Fig. 4. HMM Baseline Pipeline

To assess the performance of our approach, we compare our pipeline to both a long-only strategy in the Russell 3000 Index (the *benchmark*) and to a simple baseline 3-state HMM using weekly returns (the *HMM baseline*). A period of January 2018 to December 2021 will be used as the testing period for our strategies.

To create the baseline model, we use the Russell 3000 data from January 1992 to December 2017 as a training set. For the baseline trading strategy, we go long the Russell 3000 for bull markets, short the Russell 3000 for bear markets, and have no position for static markets. To correctly label the hidden states of the HMM, we calculate the profit using this baseline trading strategy on the training set. The state with the highest return is the bull market and lowest return is the bear market.

B. REGIME PREDICTION PIPELINE

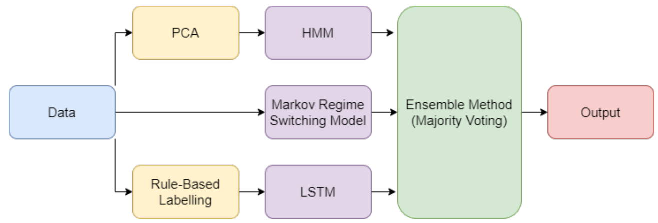

Fig. 5. Ensemble Pipeline

With reference to Fig. 5, our pipeline features four steps - data, pre-processing, modelling and ensemble. The data step (described in Section III) involves gathering the data and transforming specific series to ensure stationarity. We describe the pre-processing and model steps in the following subsections.

1) Hidden Markov Model: For pre-processing, we utilise PCA to project our data (see Table 1) down to 20-dimensions (as this explains over 90% of the variation). The low-dimensional representation of the data is the input to a three-hidden state HMM.

2) Markov Regime Switching Model: For the Markov Switching Model (MSM), we use the Russell 3000 return series as our observation series, and predict the future return for it. We choose CBOE Volatility Index (VIX) to model transition probability. We allow the MSM to fit regime-specific autoregressive models to the training data, allowing for the autoregressive, regression, and variance parameters to change across regimes.

TABLE II
PARAMETERS FOR MARKOV SWITCHING MODEL

Regimes	3
AR Lag	1
AR Switching	True
Variance Switching	True
AR Switching	True
Time-Varying Transition	True

Table II shows the values we have taken for our Markov Switching model implementation. We take 1 AR lag to make the model parsimonious while allowing for different in persistence of returns and heteroskedasticity across regimes.

3) LSTM: Our dataset consists of weekly data for the Russell index. Since there is no defined rule in the existing literature to identify bull, bear, and neutral market states we design our own methodology, which in turn help us build a supervised learning model. The initial idea came from [22] but we employ a different and simplified approach:

1) Calculate the weekly return of the underlying.
2) Take the moving average of the next 2 months of weekly returns. Bear markets are regions where the average return is less than -0.4%, neutral states have returns between -0.4% and 0.00% with all positive returns considered bull markets. These ranges were determined based on the historical distribution of returns.
3) We smooth the process by taking the mode of the next 10 signals to finalize our bull, bear, and neutral states.

With above methodology, Fig. 6 shows the classification of the regimes from 1990 to 2017 which is the period which we use to train our models.

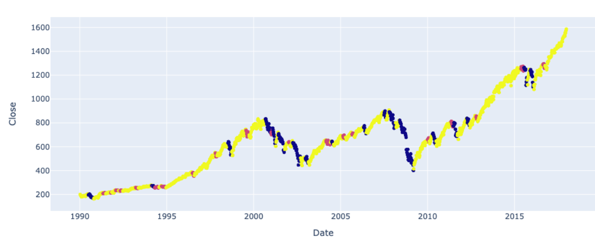

Fig. 6. Rule Based Methodology (Yellow - Bull regime, Blue - Bear regime, and Pink - Neutral regime)

To train the LSTM (with hyperparameter tuning), we divide the data into 3 parts for training (Jan 1992 - Sep 2012), validation (Oct 2012 - Dec 2017), and test data (Jan 2018 - Dec 2021). From the hyper parameter tuning, we found out the best architecture for our model which are detailed in the table IV-B.3. Figure 7 outlines the LSTM architecture used in our model.

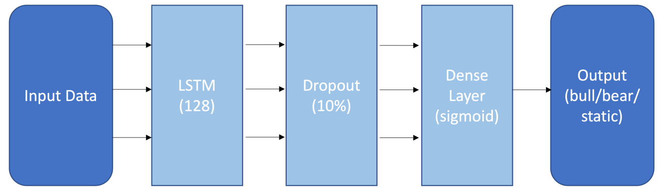

Fig. 7. LSTM Architecture

TABLE III

HYPERPARAMETERS FOR LSTM

LSTM Nodes	128
Dropout	10%
Batch Size	60
Epochs	20
Activation Function	Sigmoid
Optimizer	Adams
Loss	Categorical Cross-Entropy

Now, each of these models produce independent hidden-state predictions. To combine these predictions, we use ensemble method: **majority voting**. We take the mode of the predictions at each timestep, which thus yields our final prediction.

C. TRADING STRATEGY

We employ two trading strategies on the Russell 3000 Index to evaluate our performance. First, we test a basic trading strategy where the regimes predicted by our models are used to go either long, short or close all our position. We assume no leverage and transactions costs and, because of this, turnover rates do not impact our measured returns.

We also implement a more sophisticated version of trading strategy, taking inspiration from [1], which describes the analytical solution for the weight in the risky asset for an investor with mean-variance utility as,

$$w = \frac{1}{\gamma} \frac{\text{risk premium}}{\sigma^2},$$

where γ is the coefficient of risk aversion and σ is the variance of the Russell index. The key regime-dependent parameters in this strategy are the risk premium and coefficient of risk aversion. We use the regimes identified in our training data to estimate both. For risk aversion, we take the steady-state risk aversion as 3 [1] and use VIX in different regimes to scale this up or down,

$$\gamma_i = 3 \times \frac{\text{VIX}_{\text{avg}, i}}{\text{VIX}_{\text{avg, total}}}$$

for each regime i.

We compute the risk premium for each regime as the average of realized historical return in each of the regimes over the risk free rate,

$$\text{risk premium}_i = \frac{1}{n_i} \sum_{j=1}^{n_i} R_j - Rf_j$$

where R_j and Rf_j are the realized return and risk-free return in time period j respectively. We use VIX as a forward looking estimate of σ in the strategy.

V. RESULTS

First we show an example of the classification for the test period Jan 2018 - Dec 2021 into three different regimes by the Markov Switching Model in Figure 10. Looking at the cumulative returns for each of the hidden states, we can identify that they correspond to bull, bear and static regimes.

Also, as mentioned before, the performance of our market regime prediction model is determined by two proxy variables: cumulative profit and Sharpe ratio of a trading strategy on the test set (January 2018 to December 2021).

A. SIMPLE TRADING STRATEGY

The simple trading strategy, to recall, consisted of fully investing (long) in the Russell 3000 Index for predicted bull markets, shorting the index for predicted bear markets, and having no position for static markets. The results are shown in Fig. 8 and Table IV.

The evaluation time window is very short and is dominated by the crash in stocks that took place during the initial stages of the Covid-19 crisis in March 2020. Any model that correctly identified this episode - regardless of its longer-term sensibility or predictive power - is likely to outperform the Benchmark. Both the LSTM and MSM approaches correctly identify the bear market and are able to profit by going short. They do take some time to switch back into bull states, but cumulative returns are highly influenced by avoiding big losses and the subsequent drawdowns they experience during the market recovery are much lower than those experienced by the Benchmark and HMM-Baseline during the market crash.

Overall, the LSTM and Ensemble approaches have essentially identical results (mean return of 48.4% for LSTM vs.48.1% for Ensemble) with slightly lower drawdowns and volatility for Ensemble. These are vast improvements on the Benchmark (12.4% returns) which in turn outperforms our HMM baseline and the HMM PCA model (which is the worst performing model). The PCA model, despite its vast array of data does not react dynamically to market conditions.

B. VOLATILITY-TIMING STRATEGY

The volatility-timing strategy is a more sophisticated trading strategy with the results shown in Fig. 9 and Table V.

Here the Ensemble model performs better than the LSTM model in terms of mean annual returns (47% vs. 44%) but LTSM only experiences a drawdown of 5% vs. 11% for the Ensemble. However, the Ensemble method - because it combines information from various methods will never be the most reactive approach. It is therefore it's not surprising it doesn't have the lowest drawdown. The Benchmark outperforms the baseline and HMM PCA model again, with the HMM PCA model again having the worst performance.

Altogether, the best model in terms of Sharpe Ratio (i.e. achieving the best return with the lowest volatility) is the Vol-timing Ensemble method model. However, we emphasize again that time-window is short and differences in Sharpe Ratios not statistically meaningful. That said, the Ensemble method has a clear advantage in that is diversifies model risk. LTSM did well at identifying the Covid crash and may indeed continue to be the best performing model out-of-sample. However, there may be other situations where alternative models do better. The Ensemble approach was able to provide strong overall results even though it was incorporating information from under-performing models. This shows that

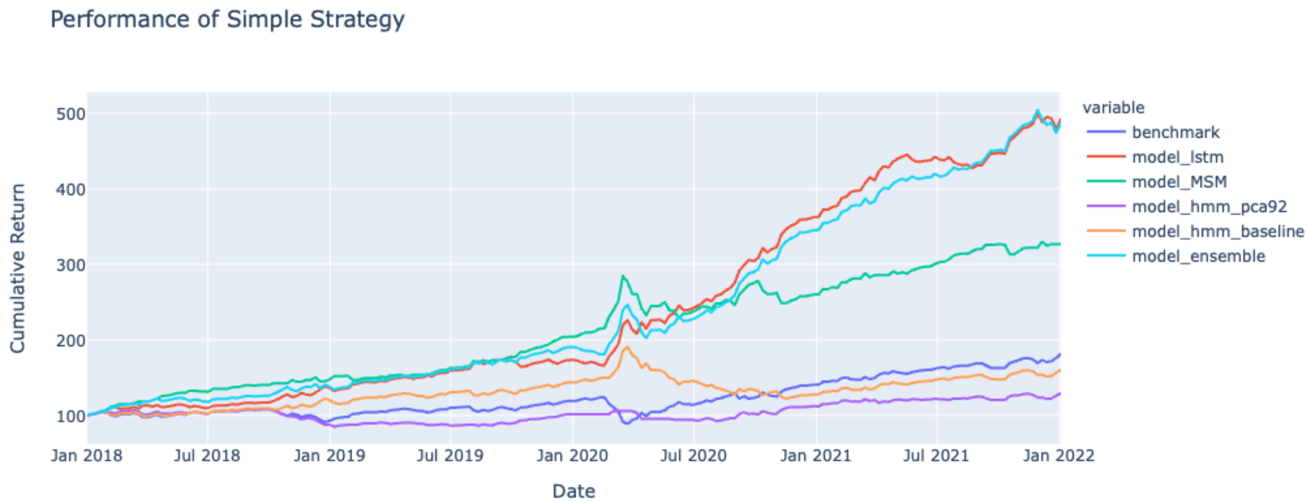

Fig. 8. Trading PnL with Simple Trading Strategy, Comparing the models to long-only strategy in the Russell 3000 Index (*benchmark*) and baseline 3-state HMM using weekly returns (*hmm_baseline*).

TABLE IV

SIMPLE STRATEGY STATISTICS

	Benchmark	HMM-Baseline	LSTM	MSM	HMM-PCA	Ensemble
Mean Return (ann.%)	15.8	12.4	48.5	34.1	6.8	48.1
Volatility (ann.%)	15.8	14.9	15.3	14.2	11.4	14.9
Skewness	-1.33	0.59	1.26	0.79	-0.18	0.86
Kurtosis	7.55	7.19	6.65	10.43	3.15	7.31
Maximum Drawdown (%)	-28.6	-36.0	-8.0	-19.6	-21.2	-17.9
Sharpe Ratio	1.0	0.8	3.2	2.4	0.6	3.2
Semi-Deviation	0.14	0.11	0.08	0.10	0.08	0.10

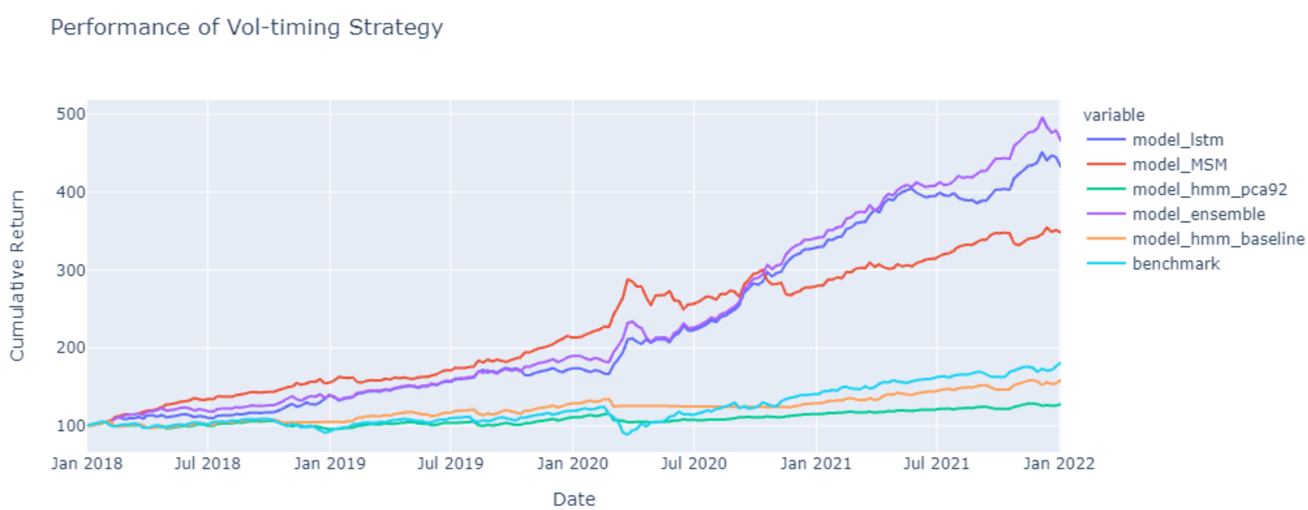

Fig. 9. Trading PnL with Vol-timing Trading Strategy, Comparing the models to long-only strategy in the Russell 3000 Index (*benchmark*) and baseline 3-state HMM using weekly returns (*hmm_baseline*).

it has a element of robustness. Further, it insulates one against the risk of relying on a single approach which may lose it's predictive power as market environments changes. It does so at the

TABLE V

VOL-TIMING STRATEGY STATISTICS

	Benchmark	HMM-Baseline	LSTM	MSM	HMM-PCA	Ensemble
Mean Return (ann.%)	15.8	12.2	44.0	36.4	6.3	46.7
Volatility (ann.%)	15.8	9.6	12.8	12.8	8.2	13.2
Skewness	-1.33	-1.13	0.66	0.16	-1.95	0.58
Kurtosis	7.55	5.47	1.84	4.09	9.10	2.69
Maximum Drawdown (%)	-28.6	-8.2	-5.4	-13.4	-10.1	-11.0
Sharpe Ratio	1.0	1.3	3.4	2.8	0.7	3.5
Semi-Deviation	0.14	0.08	0.07	0.09	0.08	0.08

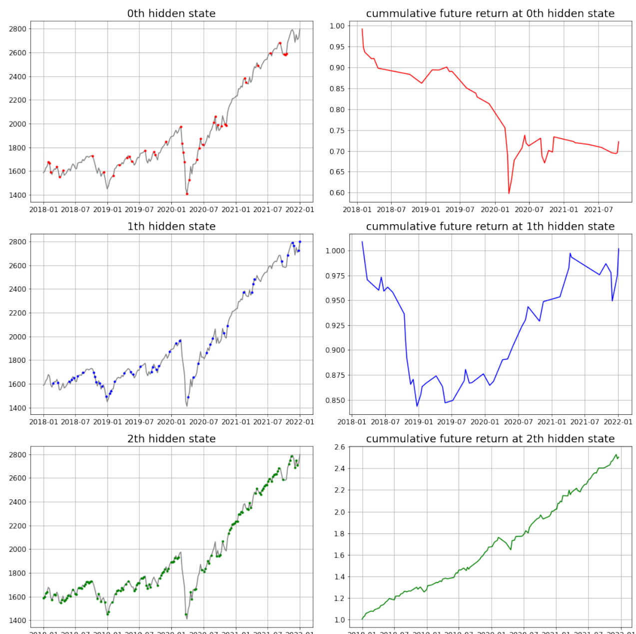

Fig. 10. Classification of Bull bear and static state by MSM model. The left columns shows the close value of Russell 3000 on the day of classification. Right columns shows the corresponding return series for a long only strategy only on a given signal. The red, blue and green plots refer to bear, static and bull regime respectively.

expense of reacting somewhat slower, but as our testing shows this may not be a large trade-off.

VI. CONCLUSIONS

From our evaluation of different techniques we conclude that Ensemble methods, for determining market states, do perform better than any single constituent model and give more robust results. Given the short time horizon to test our model, we need to be careful about drawing conclusions from the result. Any model which does well during the shock at 2020 will tend to perform better. Ensemble methods help reduce model risk as it combines intelligent constituent models while allowing room for error individually. It's very evident from the figure 8 and 9 that the Ensemble method did just that during the shock of March 2020.

Additionally, the vol-timing strategy improves the Sharpe Ratio for our models slightly, while reducing the worst case drawdown for all of them. The vol-timing strategy also allows us to come up with regimes which need not exactly map to economic {*bull*, *bear*, *static*} regimes. As the allocation weights are computed based on utility maximization, we come up with non-zero allocation during the static market regime which helps modestly improve performance.

Finally, it is important to note that the very high Sharpe Ratios that we observe are obviously never achievable in practice because of transaction costs, implementation lags and various other market frictions. Nevertheless, the trading strategies here are just a method of validating our classification of market into three regimes and underpin the utility of Ensemble methods to classify latent market states.

VII. FUTURE WORK

Using our rule-based labelling methodology enabled us to utilize supervised-learning algorithm. This opens the door for other Ensemble techniques to be explored. In particular, Boosting (training a series of classifiers wherein future classifiers focus on the mis-classification of previous classifiers) and Stacking (training a combiner model to learn to optimally combine the predictions of the other algorithms).

We can also explore adding more fundamental, macroeconomic, and survey data in the model that can work as a leading indicator to help us identifying these hidden states in the market.

For our models, we have not included any technical indicators which might be useful in capturing the behavioural patterns of market players, and can be used to time the entry and exit strategies along with our fundamental classification of regimes.

References

[1] Ang, A. (2014). Asset Management: A Systematic Approach to Factor Investing (Financial Management Association Survey and Synthesis) Chapter 2, Pg. 60 *Oxford University Press*

[2] Ang, A. & Bekaert, G. (2015). International Asset Allocation With Regime Shifts, Retrieved from `https://academic.oup.com/rfs/article-abstract/15/4/1137/1568247`

[3] Ang, A. & Chen, J. & Xing, Y. (2006). Downside Risk *The Review of Financial Studies*, 19(4), Retrieved from `https://www.jstor.org/stable/4123472`

[4] Berg-Kirkpatrick, T. & Frederking, R. (2017) Notes on Hidden Markov Model. Retrieved from `https://www.cs.cmu.edu/~tbergkir/11711fa17/recitation4_notes.pdf`

[5] Coudert, V. & Gex, M. (2008). Does risk aversion drive financial crises? Testing the predictive power of empirical indicators *Journal of Empirical Finance, 2008, vol. 15, issue 2*, Retrieved from `https://www.sciencedirect.com/science/article/abs/pii/S0927539807000679`

[6] Esmael, B., Arnaout, A., Fruhwirth, R. K. & Thonhauser, G. (2012). Improving time series classification using Hidden Markov Models, *2012 12th International Conference on Hybrid Intelligent Systems (HIS)*, 502-507, Retrieved from `https://ieeexplore.ieee.org/document/6421385`

[7] Gabriel, P. & Timmermann, A. (2000). Firm Size and Cyclical Variations in Stock Returns *The Journal of Finance*, Retrieved from `https://www.jstor.org/stable/222451`

[8] Haas, M., Mittnik, S. & Paolella, M. (2004). A New Approach to Markov-Switching GARCH Models, Retrieved from `https://academic.oup.com/jfec/article-abstract/2/4/493/900480?redirectedFrom=fulltext`

[9] Hamilton, J-D. (1989). A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle, Retrieved from `https://www.jstor.org/stable/1912559`

[10] Kim, C-J. Nelson, C-R. & Startz R., (1998). Testing for mean reversion in heteroscedastic data based on Gibbs sampling augmented randomisation, Retrieved from `https://www.sciencedirect.com/science/article/abs/pii/S0927539897000157`

[11] Filardo, A.J. (1994). Business cycle phases and their transitions, Retrieved from `https://www.researchgate.net/publication/5052311_Regime_switching_with_time-varying_transition_probabilities`

[12] Haase, F. & Neuenkirch, M. (2021). Predictability of Bull and Bear Markets: A New Look at Forecasting Stock Market Regimes (and Returns) in the Us. *CESifo Working Paper No. 8828*, Retrieved from `https://ssrn.com/abstract=3767700`

[13] Hochreiter, S. & Schmidhuber, J. (1997). Long Short-Term Memory. *Neural Comput*; 9(8): 1735–1780. Retrieved from `https://doi.org/10.1162/neco.1997.9.8.1735`

[14] Ilhan, F., Karaahmetoglu, O., Balaban, I., & Kozat, S., S. (2020). Markovian RNN: An Adaptive Time Series Prediction Network with HMM-based Switching for Nonstationary Environments. Retrieved from `https://arxiv.org/abs/2006.10119`

[15] Kole, E. & van Dijk, D. J.C. (2013). How to Identify and Forecast Bull and Bear Markets? *Paris December 2010 Finance Meeting EUROFIDAI - AFFI, ERIM Report Series Reference No. ERS-2013-016-F&A*, Retrieved from `https://ssrn.com/abstract=1695501`

[16] Maheu, J. M. & McCurdy, T. H. (2000). Identifying Bull and Bear Markets in Stock Returns, *Journal of Business & Economic Statistics*, 18:1, 100-112, Retrieved from `https://ssrn.com/abstract=146531`.

[17] Mishra. V. (2018). Comprehensive and Comparative Analysis of Neural Network, Retrieved from `https://www.researchgate.net/publication/324883736_COMPREHENSIVE_AND_COMPARATIVE_ANALYSIS_OF_NEURAL_NETWORK`

[18] Nikolaev, N. Y., Smirnov, E., Stamate, D. & Zimmer, R. (2019). A regime-switching recurrent neural network model applied to wind time series, *Applied Soft Computing*, 80, 723-734, Retrieved from `https://doi.org/10.1016/j.asoc.2019.04.009`.

[19] Nyberg, H. (2013). Predicting bear and bull stock markets with dynamic binary time series models, *Journal of Banking & Finance*, 37(9), 3351-3363, Retrieved from `https://doi.org/10.1016/j.jbankfin.2013.05.008`.

[20] Olah, C. (2015). Understanding LSTM Networks, Retrieved from `https://colah.github.io/posts/2015-08-Understanding-LSTMs/`

[21] Yu, X., Chen, Z., Xu, W. & Fu, J. (2017). Forecasting Bull and Bear Markets: Evidence from China, *Emerging Markets Finance and Trade*, 53:8, 1720-1733, Retrieved from `https://www.tandfonline.com/doi/full/10.1080/1540496X.2016.1184141`

[22] Zegadło, P. (2022). Identifying bull and bear market regimes with a robust rule-based method. *Research in International Business and Finance*, 60. Retrieved from `https://www.sciencedirect.com/science/article/pii/S0275531921002245`