GB Energy Trading with Machine Learning

Joshua Gowdy

June 12, 2025

Contents

0.1 Abstract

In this project we explore paper trading in GB day-ahead energy markets. We wish to trade the spread between daily auctions with the aid of a machine learning model to forecast these price spreads. We formulate a binary classification problem and predict the direction of the price spread using an ensemble of logistic regression and XGBoost algorithms. We define and test a trading strategy which performs well on unseen data with a annualised ROI of over 16% and a Sharpe ratio of over 8 with suggestions for further improvements.

Chapter 1

Energy Markets Preamble

The below will no doubt be a simplified view of energy markets in GB but will hopefully capture the important points to get an intuitive picture and inform power trading. A reader already familiar, or not concerned, with GB energy markets can skip this chapter.

1.1 On Energy Markets

Energy is provided from generators to suppliers to consumers. So energy contracts are also exchanged between these parties guaranteeing the flow of energy. In this study we will focus on markets which are designed to bring together generators and suppliers i.e. the purchasing of wholesale energy. It is within these markets that we wish to trade. This project does not examine the sale of energy from suppliers to consumers which anyway is not speculated on or directly traded on an exchange.

When it comes to energy trading, the actual trading is done through the sale and purchase of these energy contracts. These energy contracts act as futures requiring the delivery of energy at a specified time. Hence there are additional risks for energy traders coming in the form of costs applied to energy imbalanced positions (e.g. a trader in a position of power excess or deficit as reflected in the contracts held). One way to mitigate these risks is to be an asset-backed trader which has the ability to actually provide/use the energy that underlies the contract. If a trader is not asset-backed an imbalanced position will require balancing by the system operator possibly at a premium cost to the trader. Even asset-backed traders must sometimes must be rebalanced if forecasted generation/demand does not match actual generation/demand and so the contracts they hold may not be fulfilled. This balancing is done in a dedicated 'market' call the balancing mechanism and occurs just before the settlement period (when the power is actually provided).

Currently there is a single bidding zone for the GB wholesale energy markets and prices are uniform across GB. There is an alternative zonal pricing model where prices are independently determined in various zones based on local supply and demand. This report focuses only on the single bidding zone model. Despite being geographically a single market, there are still multiple markets for contracts exchanged on separated timescales. Namely, there are various different markets/protocols for buying or selling power contracts with the same settlement period but distinguished by how far in advance the contract is made. The day-ahead market is one such market.

Before describing the day-ahead market which will be our focus, first we note that generators and suppliers can engage in longer term contracts (with a lifetime greater than a single day) to provide power at a given price on a given date. These may be custom forward contracts made bilaterally or standardised exchange-traded energy futures with a timescale greater than a single day ahead. The purpose of such contracts is to cover in advance the baseload and minimum peak power demand to mitigate the significant portion of risk coming of (hugely) imbalanced positions.

On top of these long-term contracts there is then a need to 'shape' the demand/-supply balance closer to the time the power is provided. This is the purpose of the day-ahead and intra-day markets. The day-ahead market is important for two reasons: 1. the market is close enough to the provision of power that demand and supply can be forecasted accurately and prices will reflect these underlying fundamentals while 2. the market is not too close to the settlement period that almost all of the demand and supply has been matched already. This means there is still significant available liquidity in the day-ahead market and most trades can be met and that most traders are price takers (i.e. typical individual trades are small compare to the total traded volume). On the other hand, the intra-day markets, which enable participants to modify their position closer to gate closure, often suffer from low liquidity, meaning that it is difficult for participants to find counter-parties to trade with.

1.2 Day-ahead Markets

Day-ahead markets are typically double-blind auctions into which generators and consumers submit anonymous trades. These trades are either: an offer to generate a certain amount of power at a given price or a bid to consume a certain amount of power at a given price. Supply and demand are compared across the auction and a market clearing price is calculated for each period of the following day. The clearing price is the agreed price such that volume of accepted bids equals the volume of accepted offers and so every trade is matched on the exchange. This one price is applied to all accepted bids (those willing to pay at least than the clearing price) and all accepted offers (those will to receive the clearing price or less).

In GB their are two sequential day-ahead auctions, one in the morning and one in the afternoon. In these auctions traders place a portfolio of trades one for each of the 24 hourly periods of the following day (the second auction has the option to also place trades at a half-hourly resolution but we will ignore this possibility). The idea being that between these auctions the forecasts of generators and suppliers improve and they may wish to fine tune their position via the second auction before entering the balancing mechanism.

Generally a generator/supplier puts in offers/bids that are limited by the respective marginal cost/value of producing/supplying energy. A wind generator may place offers which are close to 0 GBP/MWh since its marginal costs are very low and may even be negative since the generator is receiving a subsidy in the form of a contract for difference. They receive a subsidy which ensures that they receive at least the guaranteed strike price of the CfD for their generation. Typically such generators are not setting the clearing price since they are low down in the merit order, they do not need to forecast the clearing price in order to place their offers. The generators which may need to forecast the clearing price directly are those that believe they may be the price setter and are in competition with each other near the bottom of the merit order. These generators have higher marginal costs set by fuel costs e.g. gas fired power plants. They need to place offers which are below the clearing price in order to be accepted, this may be just above their marginal cost or at times of scarce renewable energy far above their marginal cost. Predicting these fluctuations of the clearing price is important for these dispatchable generators to be able to cash in on 'scarcity rent' in the market.

At times of high wind and solar generation the clearing price of wholesale energy can drop to zero or even be negative, since the marginal cost of these generators is very low. For regulatory reasons when the clearing price is negative the renewable generators who hold contracts for difference are not allowed to cash in this contract and so they do not receive the strike price. It is thus important also for these generators to forecast when prices will be negative.

We will focus on non-asset backed traders trading between the two day-ahead auctions, i.e. selling to buy or buying to sell, for the underlying energy provided in each settlement period. In this case the trader must predict the clearing price in each auction in order to ascertain which strategy to follow (and possibly their expected revenue).

1.3 Balancing Mechanism

As mentioned, generators and suppliers must forecast their energy flows in advance in order to place accurate bids/offers for energy contracts to receive payment on said flows. It is common for these forecasts to not quite be right and so the trader may end up being in an imbalanced position after the day ahead (and intra-day) auctions. These imbalanced traders then entering the balancing mechanism run by NESO to ensure that come gate closure the supply and demand are matched. The balancing mechanism also allows NESO to maintain other services which are crucial for the flow of power: such as maintaining frequency of the AC current and voltage on the network. These additional services mean that the balancing mechanism is more complicated than just a very short term auction. That said a basic component of the balancing mechanism is a pay-as-bid auction for the last elements of supply and demand. Here the auction is pay-as-bid meaning that, in theory, a battery, for example, which places an offer to discharge at price X and is accepted will receive its offer price X. That said because the balancing mechanism is not just a simple auction and actually hosts assets providing other services it is often the case that said battery will be 'skipped' in favour of a higher cost generator which is able to provide and ancillary service to the grid e.g. a costly gas power plant may continue to generate in that period in order to maintain frequency/voltage. This power plant still provides power and so affects the supply/demand balance of the balancing mechanism. The battery may even be called upon to charge since the gas power plant may overshoot. Never-the-less, the cost of all these balancing actions taken ad-hoc by NESO are democratised (averaged) and passed onto the actual originally imbalanced generators/suppliers who then receive/pay a so called system price.

At times when the Net Imbalance Volume (NIV) is high i.e. there is a power deficit on the network (it is short) then the system price will be higher than the clearing price since costlier generators (possibly offering ancillary services) which were not accepted during market clearing will be called upon. When the system is long (power excess) the system price will be low in order to attract frugal suppliers in what will be a 'buyers market'.

For an energy trader in the day-ahead market, asset backed or otherwise, the system price represents a risk and an opportunity. An imbalanced trader is exposed to this system price, there is a risk in being long in a long market and being forced to sell energy at a lower price. Or being short in short market and being forced to purchase energy at a premium, this case is particularly bad since in general price swings to the upside tend to be greater (volatility smirk). This is true of most commodities and is similarly true of energy. However, traders can also position themselves favourably: for example by being long in a short market i.e. withholding uncontracted energy in order to sell it last minute at a premium or by being short in long market i.e. betting that the system price will be low and buying up that cheap energy last minute to meet lucrative contracts held from previous day-ahead trades.

Chapter 2

Problem Statement

The previous chapter contained a lot of additional background detail that is good to know but not directly relevant to building a trading strategy. In this chapter we whittle down the market specifics to a distilled form. We formulate the GB energy market from a more financial markets point of view.

For the three GB markets (morning, afternoon and balancing) we can identify three spreads which we may trade. For each spread we may either trade 'long', which by definition profits when the spread is positive, or 'short', which profits when the spread is negative. Trade combinations of more than one spread are possible but always amount to trading at most two spreads since the third is the sum of the other two. We could trade the afternoon-morning spread and balancing-afternoon spread for example. In fact given a certain amount of volume at risk it is always best to trade the largest spread, if this largest spread was known in advance. There is risk-reward trade-off here, the spread between the balancing and afternoon auction is typically larger (see Figure 3.1) but also more volatile. We opt to trade the safer afternoon-morning spread only. A more sophisticated model would be able to predict/trade both spreads and we would design and test strategies which for example either: always trade the largest predicted spread, or trade the spread the model is most confident on, or trade each individually or with some weighting. We leave this for future work.

For now we only trade the afternoon-morning spread. So we make following assumptions:

Simplifying Assumption One: We are a non-asset backed trader which always maintains a balanced position. This means that, in the day ahead auctions for every bid contract we hold we must also hold an equivalent offer contract (net balanced volume for every settlement period). This also means we are completely insulated from the risky balancing mechanism and so only trade afternoon-morning spread.

Simplifying Assumption Two: We place bids and offers such that they are always accepted. Since we are a small, paper-only trader we are not constrained to cover a marginal cost of generating energy or conversely a target price to sell energy to consumers. We simply place offers at a very high positive price and bids at a very negative price such that they are always accepted. The price we receive is the clearing price set by the market, which we assume we do not influence and furthermore is not that important to us (we care about the spread only and can make a profitable trade at any price level). Always accepted bids/offers means we have no risk of accidentally becoming imbalanced due to unaccepted contracts.

These assumptions are idealised but pretty reasonable and so may be called the beginnings of our *Strategy*.

Given the assumptions above we have the following market 'rules':

1. We have a discontinuous trading scenario, with a single 'trade' made per day.
2. Each of these daily trades is actually a portfolio of traded pairs. Each pair consists of a trade in the morning auction and a trade in the afternoon auction. We have one pair for each of tomorrow's 24 hourly settlement periods.
3. The traded pairs are for fungible energy contracts. The net position of these two trades must be volume balanced.
4. Trades are futures which are exercised the next day. So in effect they are 'one shot' trades. We spend effectively no time in the market. (This makes finding an edge/inefficiency harder.)
5. We make returns based on the price spread arbitrage between the trade pair, multiplied by their common position size.
6. In the morning-afternoon auctions, we either buy-to-sell ('long') or sell-to-buy ('short') or no trade ('hold').

The fact that the auctions occur subsequently is immaterial to us thanks to our assumptions. We must decide the daily trade before the morning auction and then we are committed and can only close the trade, and remain balanced, by placing the opposite order in the afternoon auction. Hence we should regard the 'long' and 'short' descriptors as labels only. In reality neither trade spends any time in the market. Never-the-less, price movements between auctions and between settlement periods are auto-correlated and so continuous financial indicators for price forecasting are relevant.

We must build a strategy which determines:

1. when to trade at all (as opposed to hold)
2. the direction of the spread (to either long or short trade)
3. position sizing

In the next chapter we describe the use of machine learning to predict the direction of the price spread between the two auctions.

Chapter 3

Machine Learning Auction Price Spreads

We used an ensemble of an Gradient Boosting classifier (XGBoost) and Logistic Regression classifier to predict the binary direction of the price spread between the morning and afternoon auction.

3.1 The Data

The data is gathered from Kaggle: data here. The results of this project should be regarded as a proof of concept.

The dataset consists of GB day-ahead auction data from November 2021 to September 2022 containing prices and volumes for each auction and system prices for the balancing mechanism. Figure 3.1 shows daily price curves, Figure 3.2 shows prices over the course of the entire dataset.

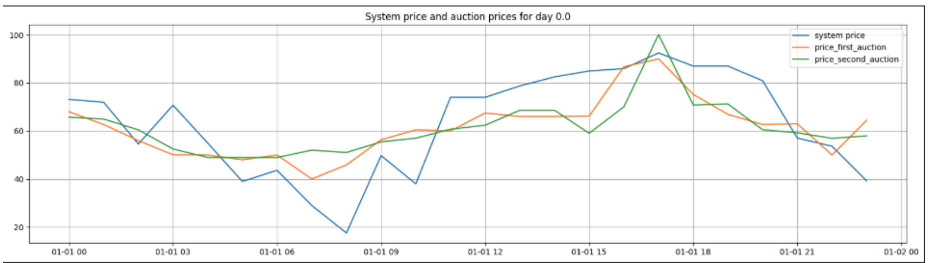

Figure 3.1: Auction and system prices. Auction prices are highly correlated and spreads are small, mean-returning and hard to predict!

Alongside this auction data we also used a collection of signals providing next-day

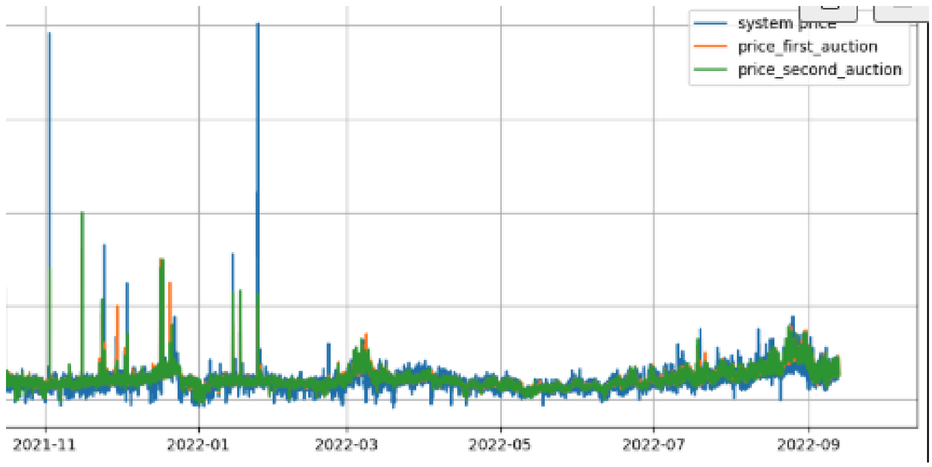

Figure 3.2: Prices move in a typical range but can spike during periods of volatility.

forecasts of energy fundamentals. These forecasts were made off-model and are not part of this project, we take them at face value. The work carried out in this project could be applied in production as part of a data pipeline which includes forecasting the energy fundamentals as a previous step in the pipeline. This previous step includes forecasting a day in advance the following energy system fundamentals:

1. supply of wind and solar on the GB energy system
2. GB power demand and supply
3. level of 'inertia' on the system
4. non-synchronous (in frequency) power due to imports

A full list of the fundamental signals can be found in the documentation folder for the project repo. All these signals can possibly have an effect on each other and on the energy prices in the auctions and the system price.

3.2 Feature Engineering - Technical Indicators

In addition to the energy fundamental signals, we also explored technical indicators to adds as features to aid in predicting the target price spread.

These included basic indicators:

1. lagged price and volume signals for the corresponding settlement period i.e. daily lags: 1 day, 3 day, 1 week.

2. close prices (i.e. the last available price from the closing settlement period of the previous day)

from which we construct more tailored indicators such as trend indicators (showing sustained price movements in a direction):

1. Rolling average percentage price change. Measures the average rising or falling of prices.
2. Moving Average Deviation (MAD): A trend energy market indicator that uses Price Moving Average (PMA) to calculate the deviation rate of the current price from PMA. Positive MAD means the price is breaking out above of its averages.
3. Z score which measures the distance from the rolling mean in units of the St Dev

volatility indicators (showing the absolute range of price movements):

1. average True Range (ATR): three different values calculated: (a) highest price minus lowest price; (b) highest price minus a 24-hour lagged electricity price; and (c) lowest price minus a 24-hour lagged electricity price. The maximum of these three values is selected for each trading hour and averaged over a rolling 24-hour window.
2. rolling St Dev

oscillator indicators (showing location of the price within its recent range):

1. Relative Strength Index (RSI): compares recent price gains to recent price losses. This indicator oscillates between 0 and 100, with a value close to 100 signifying that the majority of electricity price units within the period are 'up' and a value close to 0 signifying that the majority of electricity price units are 'down'.
2. Percentage Range (PR): An oscillator energy market indicator that finds a relationship between current electricity price and the highest/lowest prices. This indicator oscillates between 0 and 100, with a value tending towards 100 signifying that the current electricity price is closer to the lowest price and a value towards 0 signifying that the current electricity price is tending towards the highest price.

We found that including all of these indicators lead to a complex model which was prone to over-fitting with the data to hand. The best performing model actually only used the lagged and close indicators. We expect that given more data incorporating a handful of the above indicators will be beneficial to spread prediction. The most important features are in Figure

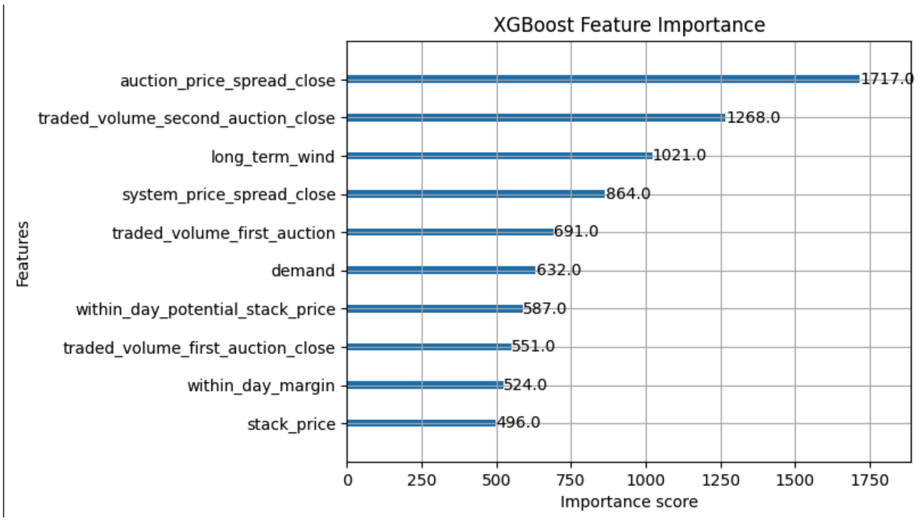

Figure 3.3: Most important of features for the XGBoost classifier.

3.3 The Model

After standard scaling the features we split the data into training, validation and holdout test sets. This is done in a manner which always preserves the chronological order of the data. The training is done using an expanding cross validation method to monitor training and check for over-fitting. This means training was carried out on an expanding fraction (fold) of the full training set and validation performed using the next consecutive fold.

The final model is that which has seen most of the data during training (80% of the data) with the final 20% split evenly into validation and holdout:

1. Holdout testing: post 2022-08-13
2. Final validation fold: 2022-07-16 to 2022-08-13
3. Final training fold: pre 2022-07-16

The validation set is used to evaluate performance and guide hyper-parameter tuning. While the holdout set is entirely unseen during modelling and represents an honest estimate of model/strategy performance.

Hyper-parameter tuning was performed using Optuna.

The final model has class probabilities determined by the weighted sum of the class probabilities from the XGBoost and logistic regression classifiers with a weighting split close to even of 47.4% given to logistic regression and 52.6% to XGBoost.

The prediction threshold is the probability above which the class is predicted to be 1 which corresponds to a positive afternoon-morning spread. The best performing threshold which lead to the highest prediction accuracy was $t = 0.462$. This threshold achieved an accuracy of $\text{acc} = 0.620$ on the validation set. This accuracy counts correct predictions equally regardless of the class prediction. It is fair to use such a naive accuracy since the class balance on all folds of the data was close to even (e.g. the overall class balance was 47.5% positive spreads and 52.5% negative spreads). Furthermore the importance of predict either spread direction is the same in both cases. We do not have any bias towards trading in either spread direction.

We also explored other measures of performance such as the ROC curve and the area underneath it. This captures the performance of the model across various thresholds and observes the respective rates of true positives and false positives (positive = class 1 = a positive spread) as the threshold to predict the positive class is relaxed. A good model predicts more true positives and false positives and so has a ROC curve above the 45 degree line, see Figure 4.2, and an area under the curve greater than 0.5 which would correspond to random guessing.

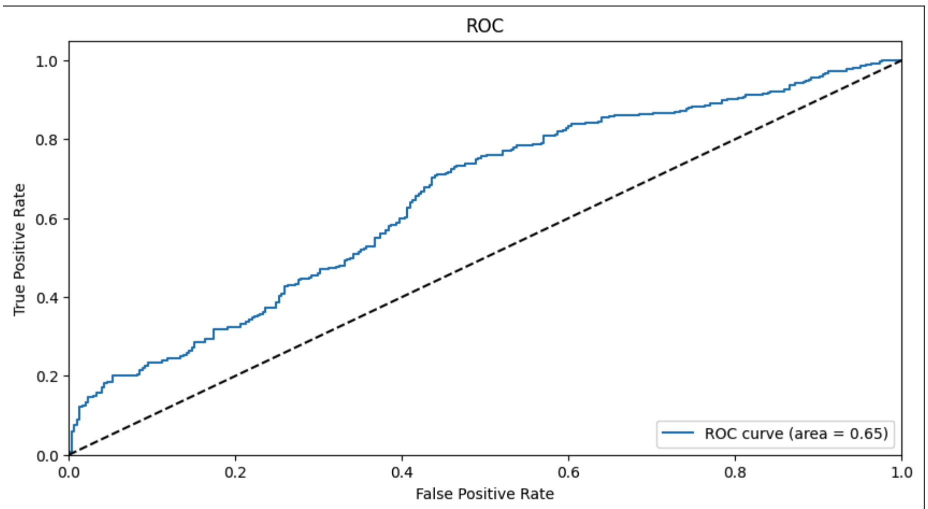

Figure 3.4: ROC curve for the model with an area under the curve of 0.651.

An model accuracy of acc = 0.620 might seem quite low but this corresponds to the model predicting a class (i.e. suggesting a trade) for every period. Instead we wish to have two thresholds one to trade long and one to trade short and for class prediction probabilities within these thresholds, i.e. close to even, the model is undecided and we hold. Following this methodology with thresholds +/- 1.4 standard deviations from the best threshold, the accuracy jumps to 73.3% while the number of trades made drops to 19.8% of all periods. There is an obvious trade off here between the number trades made and the quality of those trades as influenced by the models confidence. More trades means more capital at risk and if this leads to diminishing returns this will impact out return on investment. A balance must be struck alongside also determining position sizing. We discuss this in the next chapter within our trading strategy.

Chapter 4

Trading Strategy

We now outline our strategy to trade. We following the market rules described so far and use the the class prediction probabilities for the price spreads as our trade signal. A strategy answers the following questions:

1. when to trade at all (as opposed to hold) **(Q1)**
2. the direction of the spread (to either long or short trade) **(Q2)**
3. position sizing **(Q3)**

For all trades we assume a flat fee of 5 GBP a flat fee is a common fee model for power exchanges and is what is used on both EPEX Spot and Nord Pool exchanges.

We evaluate all strategies on the validation and holdout sets consisting of around 2 months of hourly trade opportunities none of which was seen during training and for which the only validation set was used to tune the model. A proper back/forward testing of a strategy would require a lot more data, ideally unseen, so for now we regard this as a toy example or proof of concept for each strategy.

4.1 Baseline

We first define a naive baseline strategy. This strategy uses no modelling and simply propagates the spreads observed from the previous day to the current day and trades according to those. It has a constant position size of 10 MWh **(Q3)** and never holds **(Q1)** and so trades every single period choosing short or long based on yesterdays spread **(Q2)**.

This baseline strategy had the following results after the two month period:

invested cash GBP	total profit GBP	ROI %
£4,573,678.50	£1,529.40	0.033
time held in years	annualised ROI %	mean percentage return %
0.161	0.207	1.937
std percentage return %	annualised Sharpe ratio	
58.888	0.628	

It just barely broke even over the two months even though trades were made every hour with over £4.5 million invested. This capital was at significant risk, for little to no reward, as shown but the high std of returns and a very poor Sharpe ratio.

Over the trading period there was a huge drawdown, see Figure , showing that putting significant capital into GB energy trades without a strategy is a risky proposition and one should expect to lose all ones money. Now lets try a more considered strategy.

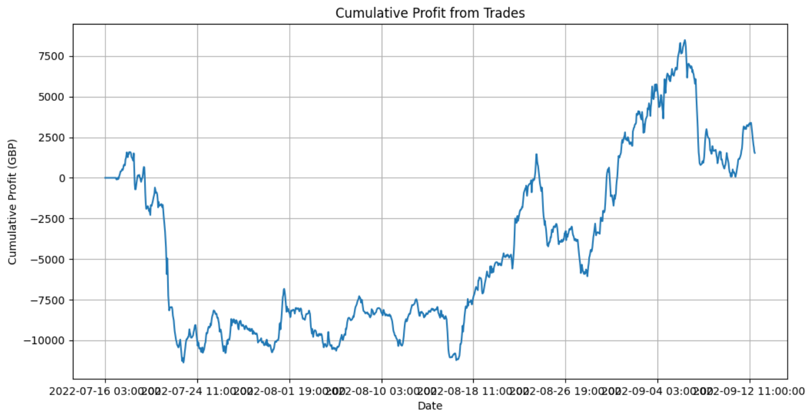

Figure 4.1: Cumulative returns of a baseline non-strategy consisting of naive copy trading. Huge drawdown observed and no realistic recovery.

4.2 The Strategy

Our actual main strategy has the following parameters with their suggested values:

1. Maximum trade volume max_vol: $V_{max} = 10$ MW. Purely illustrative and should be tailored to ones available capital.

2. Probability threshold width parameter threshold_delta: $d_t = 0.15$. This sets

the threshold for long/short at 0.15 above/below the model's class prediction threshold $t = 0.462$.

3. Position sizing_exponent: $\mu = 2$. This exponent controls how the position sizing (as a fraction of V_{max}) scales with the model's confidence. For a long trade this confidence is how close the class prediction probability is to 1 and for short it is how close to 0. An exponent of 2 means that the sizing is quite sensitive to the model confidence and periods where the model is undecided will have much smaller positions traded.

The strategy is quite simple. Call p the model's class prediction probability for a given period.

1. if $t - d_t < p < t + d_t$: **hold**
2. if $p > t + d_t$: trade **long** (i.e. buy-to-sell) with position sizing $v = V_{max} * p^{\mu}$
3. if $p < t - d_t$: trade **short** (i.e. sell-to-buy) with position sizing $v = V_{max} * (1-p)^{\mu}$

The profit per trade is equal to the spread multiplied by the volume minus the fixed fee of 5 GBP.

4.3 The Results

The strategy performs well with the suggested values of $d_t = 0.15$ and $\mu = 2$. The results for the two month period are:

invested cash GBP	total profit GBP	ROI %
£609,666.32	£15,151.89	2.485
time held in years	annualised ROI %	mean percentage return %
0.161	16.439	15.146
std percentage return %	annualised Sharpe ratio	
35.363	8.183	

Compared to the baseline, our strategy makes a healthy 2.5% ROI over the two months which gives good annual returns of 16.4%. Importantly these returns come at a low risk. Mean percentage return per trade is around +15%, with a standard deviation of approximately ±35% around this mean showing positive and consistent returns. The Sharpe ratio is a very good > 8.

The performance over time of the strategy was as follows:

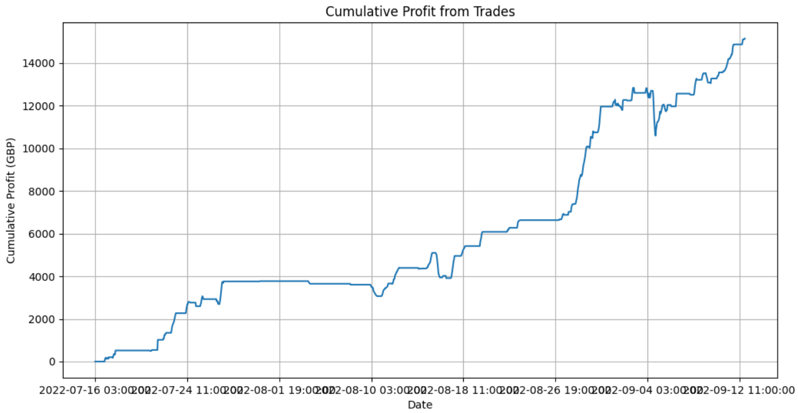

Figure 4.2: Cumulative returns of our strategy consisting of trading a long-short-hold signal from a machine learning model. Shows consistent and strong returns at least on the limited testing data.

Overall, the performance is good and the strategy merits further back-testing to see if these initial results are robust.

Chapter 5

Conclusions

We have explored paper trading in GB day-ahead energy markets. Identifying a trading opportunity which involves forecasting the price spread between the morning and afternoon day-ahead auctions and strategising risk management to insulate against the volatile balancing mechanism market. We developed a machine learning model, for a binary classification problem, to predict the direction of the price spread using an ensemble of logistic regression and XGBoost algorithms. After training and hyper-parameter tuning, this simple model avoided over-fitting and attained good enough generalisation accuracy for its predictions to be incorporated as a signal to make trades within a strategy. We develop a strategy which uses the signal both to determine when to trade and a what volume with a handful of customisable parameters. The strategy performed well on unseen data with a annualised ROI of over 16% and a Sharpe ratio of over 8. We suggest that it merits further back-testing.

5.1 Further Improvements

Apart from further back testing of the strategy to test its robustness. We also suggest some improvements which could be made:

1. Develop a regression algorithm to predict the size of the price spread not just its direction. Use this spread size to inform the volume trade. This possibility was testing during model development and it was found that a more sophisticated, possible deep learning, model was needed to predict spread size. The spreads are typically quite volatile and mean-returning on time scales which are shorter than the daily forecasting horizon.

2. Develop a strategy which allows for unbalanced positions for example trading on both the auction spread and the spread between auction and balancing mechanism. This strategy may have better returns but will likely also have higher risk if not managed properly. A successful strategy would need to determine: when to trade, which spread to trade (possibly both since the largest spread won't be known in advance), trading long or short, and finally position sizing. Broadly speaking this problem is just the doubled version of the problem we solved in this study. The final strategy will have more parameters and so will need systematic optimising.

5.2 References

I would like to thank the team at Regen for helpful discussions and advice on learning about energy markets.

1. Optimal bidding functions for renewable energies in sequential electricity markets. Benedikt Finnah. Link

2. European market integration with both physical and non-physical markets. Grigoris A. Dourbois, Dimitris Chatzigiannis, Pandelis Biskas, Anastasios G. Bakirtzis. Link

3. Risk Constrained Trading Strategies for Stochastic Generation with a Single-Price Balancing Market. Jethro Browell. Link