You can create 6 base series as follows –

1. x = xprice
2. y = yprice
3. s = spread = y – x
4. r1 = ratio1 = y/x
5. r2 = ratio2 = x/y
6. p = product = x * y

Now you can apply operators to each of these 6 series to create further features. Some examples of operators are -

1. Lag(N) – lag a series by N steps
2. EMA(N) – calculate an EMA with half-life = N
3. Diff(N) – difference between value at index (i) and index (i+N)
4. EMA(N2)-EMA(N1) – difference of two EMAs with different Ns
5. EMA(P,N) – ema of the Pth power of the series using half-life = N. P can be 2,3 or 4. This way you will calculate variance, skewness and kurtosis of the series.
6. Zscore(N) – calculated using rolling mean and stdev of last N steps

Things to note –

1. Do not look ahead i.e. any feature value at time t should only be calculated using data up to that point in time
2. Make distinction between overnight and intraday changes. This is **quite important**. Intraday change is any change calculated between two points that fall on the same day. As soon as you go over to the next day, you include the overnight market move which would be quite big. So you have to be careful to adjust for that otherwise you features would behave weirdly
    a. Let us take an example. If you are calculating diff(60) i.e. different between prices that are 60 steps apart then as your first price gets towards the end of the day, the second will go over to the next day and your difference will start including the overnight return
    b. Similar thing will happen when using EMA, lag, difference of EMAs, the moments or the z-score.

I would suggest starting with linear models. ElasticNet is a good starting point. It has Lasso in it so you will be able to perform feature selection through it, as long as you take care to account for the overnight vs. intraday effect in features correctly.