Details

Do you have to fractionally differentiate the features if one row represents one trade and they are irregularly spaced?

Yes!. In order for a classifier to learn the data generating mechanism $\mathbb{P}(\boldsymbol{x}, y)$ needs to stay consistent, i.e. the data needs to be stationary.

So even though the features are irregularly spaced, they still represent a temporal structure and thus, can have trends and seasonalities! So either do frac. diff. or integer diff. to make joint distribution $\mathbb{P}(\boldsymbol{x}, y)$ stationary.

Can I do normal CV since my labels are independent, i.e. are not based on same market information?

Yes you can do normal CV but you need to embargo it (i.e. remove immediate observations after the test set). Two reasons:

1. Lets assume fold 4 is test set and fold 5 is training set. Then some of the obs in beginning of fold 5 will be autocorrelated with the obs at the end of fold 4 (serial corr.), i.e. they are not independent from the obs in fold 4 and thus information leaks from the test set. From a trader's perspective this can be interpreted as a overlapping market regime. But this also means that some info of the test set spills into training set which we prevent by embargoing obs at beginning of fold 5.
2. closely related to previous point: if there are features in fold 5 with a look back period that starts in fold 4 than this is also an information leakage that we prevent by embargoing.

Best way to find features?

Features have to makes sense. Want more intuition in ML not less. Ideally you'd have domain knowledge and rely on tools from signal processing (wavelet transform etc.). But can also try a bunch of stuff out and do MDA approach on training, validation and test set to find best features.