COURSE MAIN GOALS:

First: To provide a demystifying review of the conceptual basis of Machine Learning through the discussion of the intuition behind the powerful and innovative ideas underlying the main algorithms: from simple regression to the perceptron, and from there to the latest neural networks with memory

Second: To provide a sound but accessible description of these algorithm's underlying mathematical structure, so that their potential innovation could be conceived and/or authoritatively discussed at the deepest level, at any forum or interview

Third: To provide fluency at the handling of the coding tools used to implement, to test rigorously -- and to adapt to the participant's needs-- the different working machine learning programs that will be provided by the instructors.

Fourth: To teach through a hands-on approach how to operate these working machine learning programs to trade different asset categories (Stocks, Currencies, Crypto Currencies, ETFs, etc.) under real market constraints

COURSE OUTLINE

Artificial Intelligence as Machine Learning, evolving from single-layer Neural Networks to Deep Learning, is fast becoming relevant to stock market forecasting. Major advances in hardware and algorithmic improvements in the last 10 years has made it possible to run deep learning models --on a laptop-- that would have been intractable before. In this course we'll give an overview of several applications of machine learning to stock market forecasting (including high frequency trading), beginning with regressions, two "shallow" machine learning models (Support Vector Machines and basic Neural Networks) and ending with a deep learning model (Long Short Term Memory Networks). Each model is discussed in detail as to what input variables and what architecture is used (rationale), how the model's learning progress is evaluated and how machine learning scientists and stock market traders evaluate the model's final performance, so that by the end of the course, the students should be able to identify the main features of a machine learning model for stock market forecasting and to evaluate if it is likely to be useful and if it is structured efficiently in terms of inputs and outputs.

COURSE OBJECTIVES

This course intends:

To present the intuitive ideas behind the most important machine learning algorithms, their underlying mathematical structure and their code implementation
To expose the student to "linear" and "kernelized support vector machines" and to provide an introduction to their use for more advanced classification tasks in finance
To introduce the student to basic neural nets: the intuition behind them and their underlying mathematical structure; to their evolution from classic and logistic regression to the perceptron; to the use these tools for time-series forecasting and classification tasks in finance.
To compare the results of time-series forecasting through examples of regressions, support vector machines and neural networks.

To show how to select data inputs for time-series forecasting by introducing machine learning data pre-processing (Data Smoothing and Data Scaling & Normalizing Techniques) and feature engineering techniques (Feature Selection, Feature Combination, Principal Component Analysis)

To introduce the main concepts related to regime indicators: "non-stationarity" of stock market returns & multiple regimes, unsupervised data clustering techniques like Gaussian mixes (continuous case), and k-means clustering

To expose the student to model evaluation and improvement: cross validation, grid search, statistical evaluation metrics and scoring, "trading-based" evaluation metrics in model selection, and "horse race-type" tests: White's reality check for data snooping, Timmerman & Pessaran and Anatolyev & Gerko tests.

To show how Deep Learning (Multiple Layer NNS, Convolutional NNS, Simple Recurrent NNS and Long Short Term Memory NNS) is used for time-series forecasting.

How to do data pre-processing, feature extraction and model evaluation and improvement without breaking the continuity of the time series.

If time allows, to introduce the basics of market microstructure theory: how a high-frequency limit order book works: iceberg orders and predictive features that can be extracted from the order book (e.g. bid-ask spread, bid-ask asymmetry etc.) to be used as inputs for Deep Neural Network algorithms.

COURSE STRUCTURE AND CONTENT:

This course consists of four parts, each one discussed in several weekly sessions:

First Part: Reviewing the theoretical basis and the intuition behind Machine Learning

(Overview of the intuitive ideas behind the basic machine learning algorithms, their underlying mathematical structure and their implementation in Python)

Second Part: General applications of machine learning to financial markets

(Simple regressions compared to support vector machines, trees & neural networks when forecasting time-series; data processing and 'feature extraction'; market model evaluation techniques)

Third Part: Deep Learning in time-series forecasting, multi-layer neural networks

Fourth Part (3 Sessions): Neural Nets with Memory and their applications to financial markets (Co-Integration & Variance Ratio as statistical metrics of autocorrelation; machine learning algorithms with memory: RNNs, LSTMs and their uses in time-series forecasting, evaluation with walk-forward methods)

--COURSE LAYOUT—

class 1

Why machine learning and deep learning are relevant to finance

Modelling for inference and modelling for prediction

Traditional statistical modelling vs machine and deep learning modelling

Types of models (by data): cross-sectional vs time-series

Types of machine learning models: supervised, unsupervised, semi-supervised, reinforcement

Random walks vs stationary processes

Classic asset price model

Homework reading review of:

Linear regression and mean squared error optimization

Logistic regression and maximum likelihood optimization

class 2
Basic machine learning workflow part 1
Use cross-sectional linear regression and logistic regression to illustrate:
Train-Test workflow
Train-Cross-Validate-Test workflow
Scaling, cross-validation and pipelines
Validation and test metrics
Parameter optimization: gridsearch
Metrics: MAE, MedAE, MSE, MSLE, MAD/MEAN-Ratio, MAPE, RSquared

Timeseries and feature engineering
TimeSeriesSplit utility
Autocorrelation features: ACF and PACF
Lags, date and time features
Visualizing feature importance

Regularization of regression
Underfitting, overfitting, good fit, unknown fit
Use of regularization to correct overfitting: L1 and L2 regularization
L2 regularization to correct "jumping coefficients" caused by multicollineality
Optimizing regularization parameters: scaling, pipelines and gridsearch
Scaling: Standardization, Min-Max, Mean normalization, Unit length scaling
Scaling in finance: linear detrending (deterministic trends), differencing (stochastic trends)

Homework:
Pandas exercise: calculating financial metrics of a trading system
Financial metrics: Annual return, CAGR, Sharpe Ratio, Maximum Draw Down, Calmar Ratio, White's Reality Check
Percent returns vs log returns and associated formulas

class 3
Basic machine learning workflow part 2
Feature reduction by selection or extraction: RFE, PCA, LDA
Use of pipelines for feature selection or extraction

More feature engineering for time series modelling
Simple autoregressive models: AR(1), MA(1), ARMA(1,1)
Timeseries decomposition example using pmdarima, lags and differencing

Homework:
Cocacola: timeseries regression model
Reading review of: F-statistic, correlation, multicollineality, statistical models for timeseries

class 4:

White's Reality Check
Time series split utility
Principal component analysis in depth Exmples:
PCA and the yield curve, factor modelling of stock portfolios (regression, with and without PCA)

Homework:
WRDS data wrangling exercise: downloading portfolio data, unstacking data, calculating portfolio metrics
Regression mechanics exercise
Readings on: PCA, cross validation types, time series cross validation

class 5:
Support vector machines
Using features or kernels to model non-linearities with support vector machines
Support vector machines parameter optimization, with scaling and pipelines
Support vector machine and feature engineering
Example: Support vector classifier for prediction of price movement

Homework:
Support vector regression model of credit default swaps
Regression mechanics exercise
Reading on metrics for binary and multi-class classification, dealing with unbalanced data.

class 6:
Trees
Relation between trees and binned features
Trees and extrapolation
Gini score and entropy score
Tree regularization
Tree visualization and feature importance
2 ways of trading a tree: extreme leaf trading and whole leaf trading
Tree parameter optimization
Example: factor modelling of stock portfolio (tree regressor)

Homework:
Reading on: why entropy works to measure information complexity.
Support vector classifier that predicts asset price return directionality: evaluation metrics in depth.

class 7:
Tree Ensembles
Ensembles in general, the law of large numbers and the binomial distribution
Bagging and random forests
Gradient boosting
Modelling correlated multiple outputs with trees or daisy chaining
The beta of a stock, length of beta lookback window, idiosyncratic volatility

Empirical asset pricing via machine learning: best predictors and models
Example: factor modelling of stock portfolio (random forest regressor, with PCA), Piotroski factor model (random forest classifier)

Homework:
Questions on trees
Personal project (40% of grade): Absorption ratio

class 8
Popular frameworks
The unreasonable effectiveness of data Keras building blocks:
Sequential model
Layers
Activation functions: sigmoid, (Leaky) ReLU, hyperbolic tangent,
Dense layer in detail
Dropout layer in detail
Loss functions: MSE, MAE, MAP, binary, categorical cross-entropy, custom
Optimization terms: pass, batch, iteration, epoch
Optimizers: Adam, SGD, others
Types of gradient descent: stochastic, mini-batch, batch
Batch size
Metrics: accuracy, etc. custom
Compile, fit, evaluate functions, validation parameters
Predict function
Visualizing the training and validation errors via history object
Callbacks
Regularization
Use of KerasClassifier and KerasRegressor wrappers for cross-validation and parameter optimization
Neural networks and scaling issues
Rules of thumb re. neural net architecture
Multivariate processing for timeseries
Walk-forward validation with gridsearch
Benchmarking a gridserach
Example: MLP classifier for price prediction with class weights, callback
Example: MLP regressor and classifier ensembles to predict bitcoin price

Homework:
Optional exercise using a Python technical indicator library called Finta 9ETF random forest model of price prediction.
Reading: Gradient Descent and Back Propagation

class 9:
Outlier identification
Autoencoders
Autoencoder and PCA comparison

Example: Credit card fraud identification
Scaling, oversampling
Supervised: Logistic Regression, Random Forest, SGBoost, Keras MLP
Unsupervised: PCA, autoencoder

Homework:
Exercise on bank failure (trees, tree-ensembles, pca and autoencoders)

class 10
Recurrent neural networks and LSTMS
Relation to ARMA models
Recurrent neural network logic: looped and unrolled
Recurrent neural network layer in Keras
The exploding/vanishing gradient problem
LSTM logic
LSTM layer in Keras
LSTM 3D inputs
Data generators

Example: LSTM applied to stock price prediction (with and without window normalization) Example: RNN, LSTM and ARIMA applied to massive data (web page views)

Homework:
Reading: VanishingGradientProblemHowToFixIt.docx, UNDERSTANDING LSTMS INTRODUCTION W ANIMATIONS & CONCLUSION

class 11
Clustering
Gaussian Mixtures
The credit cycle
Hiearchical_Risk_Parity

Example: Gaussian mixtures for price regime identification, credit cycle phase identification
Example: PCA and clustering for cointegrated pairs identification, PCA for eigen-portfolios Example: Hierarchical clustering for portfolio construction

Homework:
Exercise on gradient descent
Readings: ProbabilityDensityEstimation.pdf, ExpectationMaximizationAlgorithm.pdf, MaximumLikelihoodEstimation.pdf,
KMeansAndGaussianMixturesClustering.pdf

class 12
Technical indicators