**https://nix-united.com/blog/find-out-how-to-use-machine-learning-for-time-series-forecasting/**

**Time Series Forecasting in ML**

*Time Series*is a certain sequence of data observations that a system collects within specific periods of time — e.g., daily, monthly, or yearly. The specialized models are used to analyze the collected time-series data — describe and interpret them, as well as make certain assumptions based on shifts and odds in the collection. These shifts and odds may include the switch of trends, seasonal spikes in demand, certain repetitive changes or non-systematic shifts in usual patterns, etc.

Forcasting

* Demand
* Sales
* Stock prices
* Climate and weather
* Scientific studies
* Demographic
* Economic

**19 Time Series Forecasting Machine Learning Methods**

Some of them may even be deemed outdated by now. The so-called *legacy* time series forecasting methods either take more time and effort to implement or bring comparatively insufficient results (or both) as opposed to more recently introduced alternatives. However, they still may fit specific goals better than other approaches.

Then, there are *classical* methods which are well-tried-and-tested approaches that remain the default for most time series forecasting instances and are most widely used. The great thing is that they can be efficiently reinforced with the powers of ML to achieve much better results. *Topical*methods, on the other hand, are methods focused on particular situations and goals that fit specific forecasting scenarios.

As in any case of working with machine learning, ML-forecasting can be supervised (requiring specific data input to work) or unsupervised (ongoing, self-learning data processing mechanisms). The methods listed below can be of both natures interchangeably.

**Legacy Methods of Time-Series Forecasting:**

**Recurrent Neural Network (RNN)**

RNNs process a time series step-by-step, maintaining an internal state from time-step to time-step. Neural networks are great in this application as they can learn the temporal dependence from the given data. And considering input sequences from the temporal perspective opens horizons for more precise predictions. However, the method is considered *a legacy* because the “education” of neural networks can be too time-consuming.

**Long Short-Term Memory (LSTM)**

It’s kind of RNN, but while maintaining RNN’s ability to learn the temporal dynamics of sequential data, LSTM can furthermore handle the vanishing and exploding gradients problem. Thus, complex multivariate data sequences can be accurately modeled, and the need to establish pre-specified time windows (which solves many tasks that feed-forward networks cannot solve). The downside of overly time-consuming supervised learning, however, remains.

**Classic Methods of Time-Series Forecasting:**

* **Multi-Layer Perceptron (MLP)**
* **ARIMA**
* **Bayesian Neural Network (BNN)**
* **CART Regression Trees (CART)**
* **Support Vector Regression (SVR)**
* **Gaussian Processes (GP)**
* **Radial Basis Functions Neural Network (RBFNN)**
* **K-Nearest Negihbor Regression Neural Network (KNN)**
* **Kernel Regression or Generalized Regression Neural Network (GRNN)**

**Multi-Layer Perceptron (MLP)**

Univariate models can be used to model univariate time series forecasting problems. Multivariate MLP models use multivariate data where there is more than one observation for each time step. Then, there are multistep MLP models — there is little difference to the MLP model in predicting a vector output that represents different output variables or a vector output that represents multiple time steps of one variable. This is a very widely used method that even outperforms LSTM in certain autoregression instances.

**ARIMA**

Autoregression employs observations collected during previous time steps as input data for making regression equations that help predict the value to be generated at the next time step. ARIMA or an AutoRegressive Integrated Moving Average model combines autoregression and moving average principles, making forecasts correspond to linear combinations of past variable values and forecast errors, being one of the most popular approaches due to that.

**Bayesian Neural Network (BNN)**

BNN models involve constructing a prior distribution and updating this distribution by conditioning on the actual data. This is particularly useful for financial data because of its volatile nature, as nonlinear time series forecasting with machine learning is enabled. BNN treats network weights or parameters as random variables, being among the most universally used models out there.

**Radial Basis Functions Neural Network (RBFNN)**

RBF Neural Network is based on the function approximation theory or supervised and unsupervised manner was used together. Similar to BNN, RBF models are used for forecasting nonlinear time series. RBFNN model proves to be best for predicting daily network traffic, which makes it pretty popular among commercial forecasting applications.

**Kernel regression or Generalized Regression Neural Network (GRNN)**

Generalized regression neural network (GRNN) is a branch of the RBF neural network. Recent research activities in forecasting with GRNN suggest that GRNN can be a promising alternative to the traditional time series model. It has shown great ability in modeling and forecasting nonlinear time series, and it is gradually entering the lines of multipurpose, commonly used methods.

**K-Nearest Neighbor Regression Neural Network (KNN)**

The k-nearest neighbor (k-NN) algorithm is one of the most popular non-parametric approaches used for classification, and it has been extended to regression. KNN is a supervised machine learning method that consists of instances, features, and targets components. The selection of the number of neighbors and feature selection is a daunting task. KNN is a simple algorithm that has been effectively used in various research areas such as financial modeling, image interpolation, and visual recognition.

**CART Regression Trees (CART)**

The technique is aimed at producing rules that predict the value of an outcome (target) variable from known values of predictor (explanatory) variables. They show good prediction accuracy performance, but they cannot detect and adapt to change or concept drift well. This approach is certainly strong in terms of unsupervised practices, but it still lacks maturity.

**Support Vector Regression (SVR)**

(SVM) is a supervised machine learning algorithm that can be used for both classification or regression challenges. The ability of SVM to solve nonlinear regression estimation problems makes SVM quite successful in time series forecasting.

**Gaussian Processes (GP)**

Gaussian process (GP), as one of the cornerstones of Bayesian non-parametric methods, has received extensive attention in machine learning time series prediction. GP has intrinsic advantages in data modeling, given its construction in the framework of Bayesian hierarchical modeling and no requirement for a priori information of function forms in Bayesian reference.

**Topical Methods of Time Series Forecasting:**

* **Convolutional Neural Network (CNN)**
* **Attention Mechanism**
* **AdaBoost**
* **Kaggle**
* **Transformer Neural Network**
* **Decision Trees**
* **LightgBM**
* **XGBoost**

**Convolutional Neural Network (CNN)**

Although analysis of image datasets is considered their main field of application, convolutional neural networks can show even better results than RNNs in time series prediction cases involving other types of spatial data. For one thing, they learn faster, boosting the overall data processing performance. However, CNN’s can also be joined with RNNs to get the best of both worlds — i.e., a CNN easily recognizes spatial data and passes it to RNN for temporal data storing.

**Attention Mechanism**

This is one of the basic principles of deep learning that can be adapted in terms of different forecasting models. In a nutshell, it mimics the human brain in terms of focusing attention on specific elements that stand out from a bunch. This enables a deep neural network to concentrate only on relevant data points among the barrage of various inputs, boosting the efficiency of NLP and Computer Vision.

**Transformer Neural Networks**

A transformer neural network is an advanced architecture focused on solving sequence-to-sequence tasks. Its main goal is also to easily handle long-range dependencies. Such networks are quite popular in ML-based models, simplifying regression by simply customizing the *loss function*. This comes in more than handy when it comes to regressions.

**Kaggle**

Kaggle is a coding and data processing environment where efficient web traffic time series forecasting can be carried out. This is an engine with technical capabilities contributed by an extensive community of enthusiasts over the years. This makes it an efficient tool for tackling the issue of predicting future values of multiple time series.

**LightGBM**

This one is a widely used ML algorithm that is mostly focused on capturing complex patterns within tabular datasets. This results in quite efficient sales data predictions. In certain instances, LightGBM outruns the classical ARIMA approach in terms of making tabular-based predictions. However, both should be applied in individual situations to make out the best.

**Decision Trees**

ML-based decision trees are used to classify items (products) in the database. Generated classes get dedicated multivariate time series models that help predict the future price of a certain item. Obviously, this one is best for commercial analyses.

**XGBoost**

This is the applied machine learning algorithm that works with tabular and structured data. In its core, lie gradient-boosted decision trees. Working with XGBoost requires one to transform time series datasets into supervised learning problems. But the results should be worth it.

**AdaBoost**

This type of forecasting algorithm is deemed as the best out-of-the-box classifier by many. This means that it is best used at elaborating data classifications in conjunction with other efficient algorithms. For instance, when used with decision trees, it learns to outline the hardest-to-classify data instances over time.

**Step-by-Step Process of Time Series Forecasting Using Machine Learning**

**Preparation Stage**

* Project goal definition
* Data gathering and exploration, Data visualization, charts, plots
* Data preparation and time series decomposition — Clean data.

**Modeling Stage**

* Forecasting models evaluation —different forecasting models are tested and evaluated to pick the most efficient one(s).
* Forecasting model training and performance estimation — the picked machine learning algorithms for time series are then optimized through cross-validation and trained.

**Testing Stage**

* Forecasting models run on testing data with known results — a step necessary for making sure the picked algorithms do their work properly.
* Accuracy and performance optimization — the last phase of polishing up algorithms to achieve the best forecasting speed and accuracy.

**Deployment Stage**

* Data transformation and visualization — to integrate the resulting forecasting model(s) with the production at hand, the gathered data must be conveniently transformed and visualized for further processing.
* Forecasting models revision and improvements — time series forecasting is always iterative, meaning that multiple ongoing revisions and optimizations must be implemented to continuously improve the forecasting performance.

**Key Challenges of Forecasting Time Series with Machine Learning Models**

Keep in mind that you may (and will) come across certain common challenges of using machine learning for time series in the process. These include the following.

* Not enough data, or seasonal data
* Not so accurate
* Not clear how the business is run