

Introduction to Blue Yonder's Basis Technology

Dr. Alexander Beck¹, Prof. Dr. Michael Feindt¹

¹Blue Yonder GmbH, Karlsruher Str. 88, 76139 Karlsruhe, Germany
alexander.beck@blue-yonder.com

1 INTRODUCTION

Fully automated and self-learning Predictive Analytics technologies create new possibilities for creating forecasts based on large data volumes. Blue Yonder, with its Predictive Analytics Suite, provides a comprehensive end-to-end solution for challenges in this area. This technical white paper describes the structure and functioning principle of the Blue Yonder basis technology.

Overview

- **Fully automated end-to-end solution**

Blue Yonder begins at the data source and ends with the delivery of forecasts in pre-defined time intervals or on-demand in any format that one chooses.

- **Robust data processing of the highest quality**

Before modern machine learning algorithms are used, the raw data are prepared and structured through a chain of scientifically proven methods and then smoothed out (this is known as pre-processing). This enables the highest forecast quality standard to be attained. It also makes creating forecasts robust with respect to outliers and errors, as well as missing data and entries.

- **Self-adaptive**

The Blue Yonder Predictive Analytics Suite recognizes when new information (data increments) needs to be included in the way the model is set up. It does so on its own.

- **Individual probability density**

The algorithm uses Bayesian statistics and chained classifications to generate a probability density for individual events. This enables decision-making that is well grounded and that takes risk into account. In addition, cost and sales functions connected to the decision-making process can be optimized.

2 PREDICTIVE ANALYTICS

Due to the continually rising number of available data sources of diverse quality and all kinds of complexity, traditional analysis models and forecasting models soon run up against their own limits. They were neither developed for heterogeneous and continually growing data quantities, nor can they provide acceptable results in short time horizons, nor adapt quickly to trends. However, the available quantity of data allows one to distance oneself from the rigid theories and model assumptions and instead to use self-learning processes that automatically penetrate the available data structure. Whereas in the past, proprietary developments and tool solutions were the backbone of quantitatively oriented enterprise areas, now, self-learning methods and fully automated systems are the focus of attention. Highly efficient machine learning algorithms are increasingly used where complex decisions must be made

with the highest quality requirements, and in fast sequence. This development is driven by the desire to turn the reality depicted in the complex data quantities into things useful for business.

Introduction to and functionality of the Blue Yonder basis technology

The Blue Yonder basis technology is a comprehensive end-to-end solution for forecasting future events and is based on making statistical insights from large data volumes. The algorithm was first developed at the international research center, CERN, for questions concerning particle physics¹. Based on the flexibility of this technology for forecast creation for almost all time horizons, the algorithm has been used successfully in business and industrial areas for some years now. Because the Blue Yonder forecast technology is used by scientists and continues to be developed by them, the methods used correspond to the requirements needed to also prove their worth in scientific circles and to gain acceptance there.

Application areas The Blue Yonder Predictive Analytics Suite finds ideal application uses in situations in which, based on large data volumes (as of approx. 10^4 data records) forecasts are required in regular intervals (for example, per second, daily or weekly), and which need to be provided in a fully automatic manner to downstream processes, or to be manually processed downstream. The algorithm that it is based on is specially developed for problems such as those depicted in Figure 1.

Type of forecast Both simple classification problems (yes-no questions) can be solved using Blue Yonder technology, as well as requirements in which numerical values are forecast (regressions). In the latter, Blue Yonder has particular strengths.

Derivation of an individual probability density The Blue Yonder technology creates a probability density on the level of individual events, based on Bayesian statistics and a chained use of classifications.

Examples

- An individual insurance policy holder (insured), for which the expected loss sum is forecast for a period of two years.
- A cosmetics article in a specific branch of a pharmacy chain, for which a forecast is created for a specific day, with the corresponding forecast weather conditions.

The comparison between the probability densities for individual and average events is depicted in Figure 2. An average event, for

¹ A Neural Bayesian Estimator for Conditional Probability Densities, Michael Feindt, 2004, <http://arxiv.org/abs/physics/0402093>

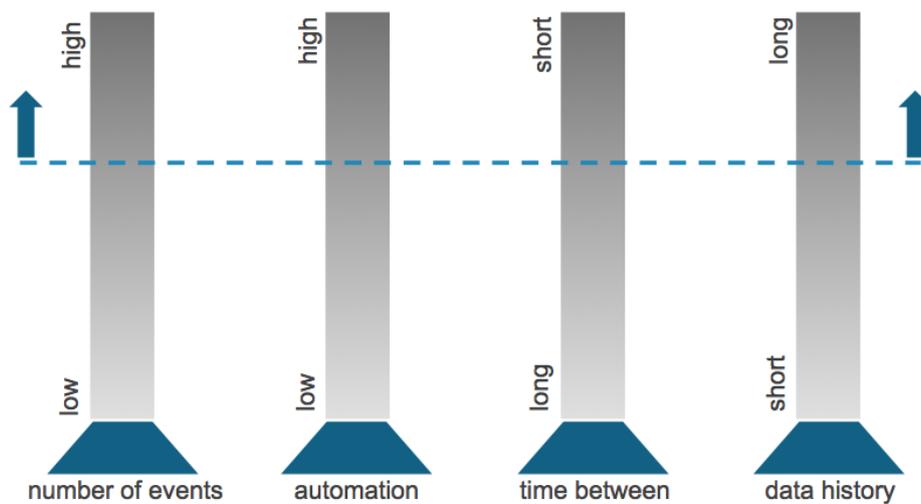


Fig. 1. Criteria for the successful use of the Blue Yonder Predictive Analytics Suite.

example, relates to the purchasing behavior of a group of people who live in large cities. An individual event in this context could be a person with the following characteristics: *Hannover, Germany as the residence, 53 years old, female, ... drivers license: Class B*. An individual forecast and view, as explained in this example, allows a considerably more precise conclusion about future events, in comparison with the statistical treatment of groups of people.

Based on the probability density, any forecast measure or dispersion parameter can be generated. It is thus possible to construct the median or a p-percent quantile. For example, the latter offers commercial enterprises the possibility of attaining a general goods availability of 95 percent. The calculation of the dispersion measure allows information on the statistical uncertainty of a forecast measure. The forecasted probability density distribution allows the construction of asymmetric errors for positive or negative divergences from the forecasted value.

This information can be used to optimize decisions with regard to the risk aspects. In this way, an optimization of the forecast measure based on downstream cost functions of any chosen kind is possible. For commercial enterprises, for example, costs that result from a goods surplus are different in their nature from costs that result from a goods shortage or bottleneck. With this technology, not only the expected sales quantity can be forecasted, but also the resulting costs that arise from divergences from the forecast value. By optimally choosing the location parameter, the costs can be reduced to a minimum.

Adjustment to data increments The Blue Yonder technology can react to data increments and include and use new information correspondingly to generate forecasts, at any time. This is particularly important for the early recognition of trends and particularities, which otherwise would only be seen in ex-post analyses. Dependent on the problem, in pre-defined intervals or using the definition of specific events, a fully automated adjustment of the model can be made. It is also possible to archive models that are no longer used, so that a full reproduction of the forecast creation using any time period is possible, at any time.

Blue Yonder Technology

In this section, the way that the Blue Yonder technology works is schematically explained. Figure 3 shows a possible connection of individual technology components. In particular, the section above pre-processing is put together and configured for each data problem, in an appropriate way.

After an understanding of the data problem has occurred, where for example the data quality and problem-specific structures are recognized, the quantity of data is then enriched by adding additional information and sources.

Examples of additional data sources comprise:

- Weather data
- Online product price information available
- Social media data and the mood indicators resulting from them
- Economic and business indicators, stock market values, mood barometers
- Digital image data
- ...

In principle, each individual data source can be used, so long as it fulfills specific criteria with respect to availability, quality and information frequency.

Based on this data foundation, the record of pre-defined, independent variables (input variables) is expanded to include additional, constructed parameters. For example, patterns that recur over time are recognized and handled separately. Classic time sequence effects such as a regular fluctuation of a volatility are modeled in this phase and made available to the downstream process steps as enhancing input variables.

Pre-processing Pre-processing is a fixed component of the Blue Yonder basis technology and is made up of a chain of process steps that prepare and smooth any type of data. In particular for

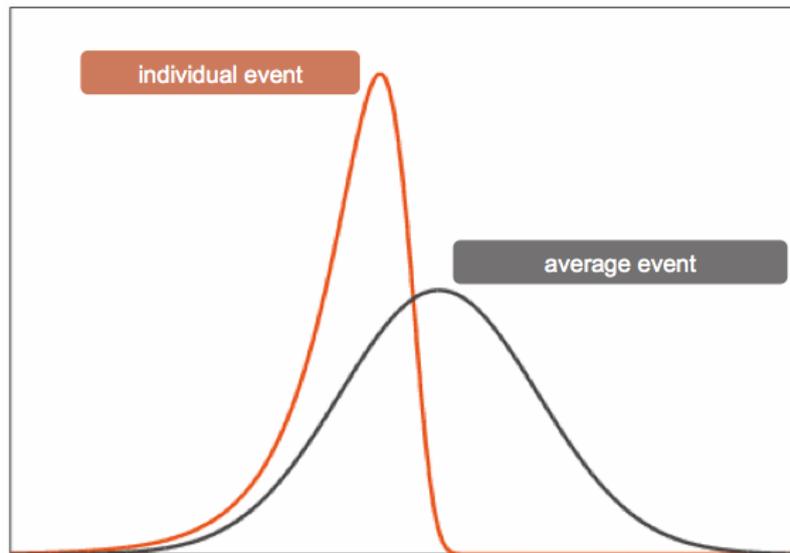


Fig. 2. The probability density of a quantity of observations and the probability density for an individual event. The quantity distribution is considerably broader than that of a single individual event. Multiple events that resemble one specific event in their nature follow the same individual probability density.

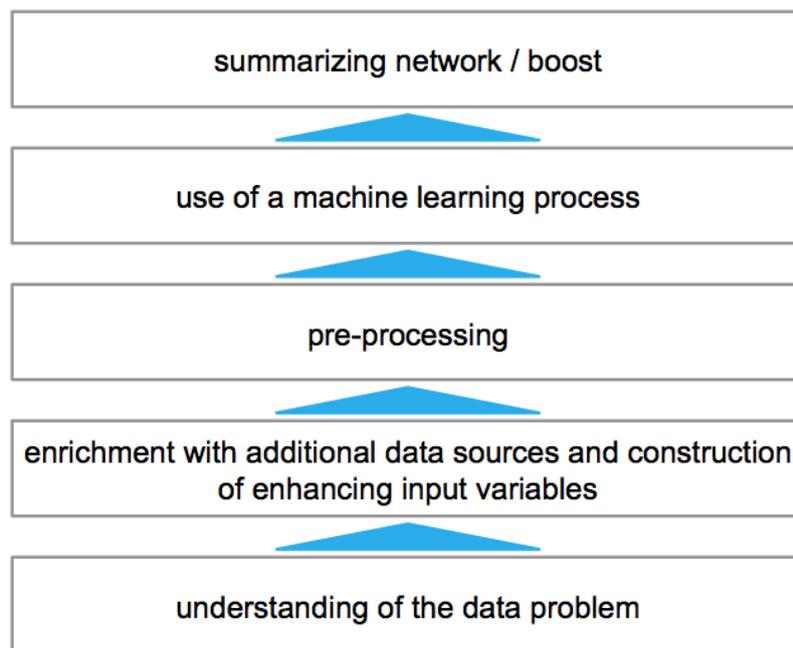


Fig. 3. Schematic representation of one possible form of the Blue Yonder core algorithm.

optimization problems in high-dimension solution spaces, focusing on this step is essential. Optimization algorithms put on a pre-structured solution space with considerably reduced computer effort required show considerably better results.

In the core of pre-processing are methods for handling missing values and outliers (outliers are values that have an incredibly high divergence from the rest of the value quantity). Over and beyond this type of error handling, regularization steps are carried out for handling very rare events that cannot be ignored. The most

prominent example here is an automobile insurance company that has only insured a single model of a specific auto manufacturer, alongside its usual pool of insured vehicles. If that model has a loss, the frequentistic loss rate for that vehicle will be an unrealistic 100 percent. A component of the forecast technology is made up of Bayesian regulation methods that use effects of this kind in a sensible and mathematically correct manner. The handling of multi-dimensional correlations is an additional example of the power of pre-processing. In this way, a variable V_1 has an exactly inverse

correlation to the target value for men and women. If this effect remained unrecognized, variable V_1 would be incorrectly calculated and would be unused. In addition, the pre-processing step is very important to guarantee the robustness of a Predictive Analytics application, including in productive operation. This applies to how data errors are handled, as well as to avoid overfitting (overfitting means that statistical fluctuations are learned).

Modern algorithms After pre-processing, modern machine learning processes are then used, depending on the problem, which can also detect and process non-linear effects and relationships. A choice of robust algorithms is available for that, which can be used individually or in combination, in order to guarantee a maximum generalization of the solution, while at the same time retaining the detail accuracy.

In addition, a chaining of algorithms or an iterative training model is conceivable, in which previously unrecognized relationships are incrementally captured on several levels and included in the model (boosting).

Introduction and operation Experienced data scientists at Blue Yonder assist with the above-described processes to create a specific forecast solution. At the end of an implementation project, there is an end-to-end solution, which is coordinated from data capture to delivery of the forecasts and which is optimized to the specific problem. During running operations, the connected data sources are automatically called on. On this basis, forecasts are delivered in pre-defined regularities or on demand, which can, on the customer side, be fully automated or put on an end-user device.

ABOUT BLUE YONDER

Blue Yonder is the leading software provider for Predictive Analytics in Europe. Blue Yonder makes a decisive contribution to enterprise success with its accurate forecasts, in real time and based on the cloud. Blue Yonder recognizes previously unrecognized connections and patterns in structured and unstructured data. The software was first developed in experimental particle physics. At research institutes like CERN, every second, the world's largest particle colliders analyze data in the petabyte range. Today, diverse industrial and business sectors use that power in diverse applications. The self-learning (machine learning) and dynamic software provides accurate sales forecasts and automatic order proposals (auto materials planning). It is successfully used in dynamic pricing as well as in customer data analysis. Enterprises are thus in position to address customers with individual product recommendations and offers, as well as to recognize customers who are about to leave, at an early time. Precise forecasts enable an exact judgment on the credit rating of existing and new customers. In addition, the analysts at Blue Yonder provide a real added value for the risk management of banks. The company was founded in 2008 and its customers include OTTO, Vodafone, and EOS. Blue Yonder has already won the prestigious Data Mining Cup three times. Blue Yonder also won the 2013 FOCUS Digital Star Award, the 2011/12 CyberChampions Award, and the 2012 CyberOne Award. Forecasts that continually and automatically adapt to changed conditions make a decisive contribution to making business processes profitable and ready for the future.