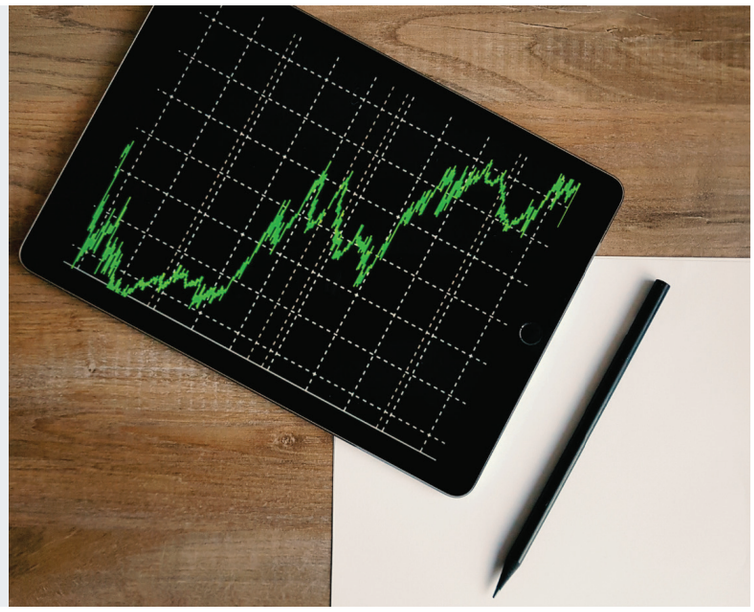
This Innovation Note was written by our Partner Knoldus

Building a Guided Analytics Forecasting Platform with KNIME



Maintaining inventory and ensuring that stock is consumed efficiently is a key decision that many companies - particularly those in retail - have to make. An excess or shortage has a major effect on profitability and can cost retailers worldwide up to $1.1 trillion annually. Overstocking can lead to decisions like marking down the item's price, which increases sales turnover. Having limited stock results in lost sales and dissatisfied customers who then purchase from the competition.

Forecasting is a basic procedure for any business, particularly those in Consumer Packaged Goods (CPG). Stock, production, storage, delivery, and showcase are all influenced by accurate forecasting. However, an accurate forecasting model may not be everything that an organization wants. It may want to involve different stakeholders in the workflow. This is where the Knoldus Forecasting Platform (KFP), which is built using KNIME, comes into play.

With the KFP, data scientists create a model to forecast sales and tune it for accuracy. Decision makers then set parameters for the forecast based on their needs. The KFP is deployed on KNIME Server as a web application via the KNIME WebPortal. This makes it an easy to use tool for users who aren't data science experts. Because even without the technical details of how the forecasting works, they can customize the input and model parameters and visualize the result of each manipulation.

Challenges

Using forecasting solutions to predict sales or stock consumption is not new. However, most organizations experience the following challenges:

- **History is not enough to predict the future.** Most forecasting systems are built on the assumption that historical data is enough to predict the future. However, with the increasing complexity of supply chains, extraneous data makes a big impact on the future. Enterprises have rigid forecasting processes, which make it impossible to implement changes and build new models. This is more acute if the forecasting is done using packaged applications because the integration of external data is complex and time-consuming.
- **No two products are alike**. Enterprises rely on demand planners who use ERP systems extensively. They know that no two regions or two products are the same. Yet the models built in the ERP are rigid and use the same model across all product attributes. This may simplify the forecasting process, but the forecast is of no use. As a result, stakeholders who have low confidence make manual changes to reflect their impressions. A better approach is to have different models for different products or other classifications. However, current forecasting data flows are complex and too intertwined to accomplish this.
- **(Un)-availability of accurate transactional data.** Transactional data is continuous, dynamic, and constantly changing. Forecasting systems are usually either embedded into large ERP applications or run on specialized statistical platforms. The difficulty of converting real-time transactional data into these systems is long and complex. Furthermore, these pipelines are developed on other software packages or custom scripts, making it extremely difficult to change or improve.

Solution

To aid in solving these problems Knoldus built the Knoldus Forecasting Platform (KFP), a web application built using KNIME that allows decision makers and stakeholders to be as equally involved as data engineers and data scientists in creating a pipeline.

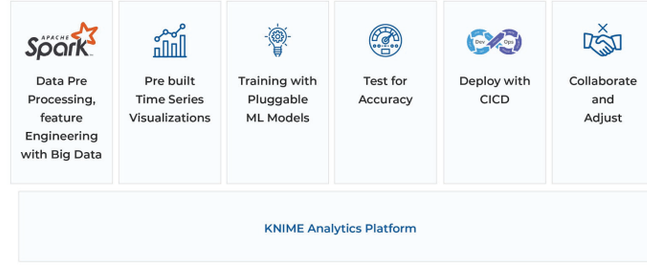


Fig 1: Overview of the Knoldus Forecasting Platform (KFP) supplied by Knoldus..

The KFP provides several advantages over historical forecasting solutions:

1. Configurable, dynamic platform. Allows the underlying forecasting process to be customized by changing the parameters, datasets, or models, which can be done within a few hours or minutes to provide a timely forecast.

2. Faster, flexible processing with big data. End to end pipelines can be run, in most cases, multiple times a day and are only limited by the computational spending users are prepared to incur. Companies can choose to generate forecasts on any cadence that they want or need.

Open for Innovation KNIME

KNIME AG · Hardturmstrasse 66 · 8005 Zurich, Switzerland · info@knime.com · www.knime.com
©2021 KNIME AG. All rights reserved.
The KNIME® trademark and logo and OPEN FOR INNOVATION® trademark are used by KNIME AG under license from KNIME GmbH, and are registered in the United States and the European Union.

Try it out for yourself!

This workflow is available on the KNIME Hub:
tinyurl.com/forecasting-platform

3. Rich set of prediction models. Machine learning allows for a quick change in models that fit what companies are trying to forecast. The biggest strength of KNIME and, as a result, the KFP, is the ability to plug in advanced models such as neural networks and random forest algorithms with no code (but coding is possible when needed), making the forecast sophisticated and accurate without increasing the complexity.

4. Accuracy measurement. Enables measuring the accuracy of the forecast following the principles of machine learning systems. Machine learning algorithms inherently come with accuracy measurements, versions of data-like training, test and production datasets, and give valuable feedback early on.

5. Ability to react to black swan events. Reduces the risk of missing out on key global events by allowing for quick changes. Many companies miss out on critical events due to the lack of an easy way to integrate external events.

6. Discipline due to implemented forecasting process. Forecasting processes are generally very well established and too rigid to change. For a well-tuned supply chain flexibility is needed to incorporate stakeholder feedback, configure different forecasting parameters, and integrate it into a legacy system. The KFP can be managed independently and integrated into existing business processes.

Workflow Steps

There are four steps to creating the KPF using KNIME Software:

1. Take a dataset in any form, load it in it's database, and offer different options to end users for data filtering and preprocessing.
2. Display end users an analytics dashboard for reading different aspects of the data.
3. Display end users a data inspection plot for seasonality, trend, and stationarity of the data.
4. Send end users all forecasting results and trained models.

Once the reports and visualizations are generated, data scientists, business users, and domain experts can collaborate on the final results.

Results:

With this Guided Analytics application, companies can create data visualization dashboards and forecasting models as well as generate forecasting for their business intelligently and collaboratively by:

- Ingesting data from different data files
- Configuring parameters for the forecasting process
- Creating data visualization dashboards in an easy and guided way
- Using already available statistical, machine learning, and AI-based algorithms
- Using an in-built email service for collaborating on results

Why KNIME Software

The free and open source KNIME Analytics Platform made the development, access, and management of this solution very easy due to the seamless integration with other technologies. For example the JavaScript Extension and Python Integration made visualization easier in the workflow itself. The wide range of customizable KNIME core nodes for data transformation helped in making tedious pre-processing and data cleaning tasks such as changing the data's structure, extracting date and time, and combining columns much simpler without needing to manipulate code. On top of the core transformation nodes, KNIME also provides nodes to remove stationarity, inspect seasonality, and perform other tasks specific to time series data. Finally, the machine learning nodes provided assistance in training different models to compare their accuracy for best performance. The vast selection of nodes, plus the procedures that are possible, make it easy to create solutions such as this. Once the solution was ready, it was simple to deploy it to the KNIME WebPortal via KNIME Server to create a powerful web application. This enabled domain experts and decision makers to become part of the process and interact with parts of the workflow relevant to them and their expertise.

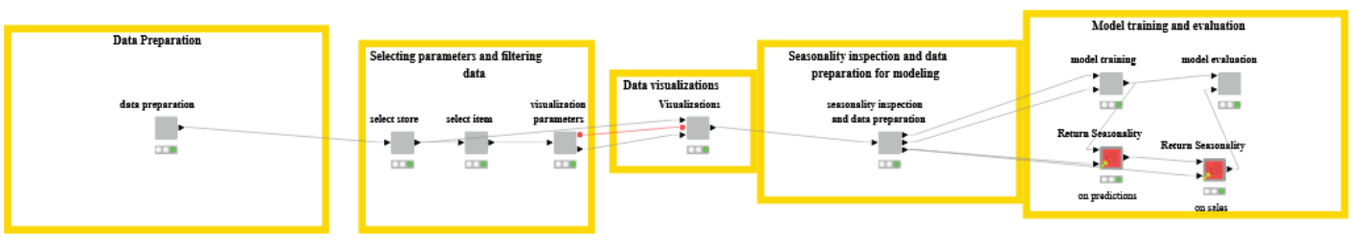


Fig. 2: Overview of KNIME workflow.

Knoldus, a KNIME partner, keeps your business competitive and future-ready with extremely well-engineered systems through the unwavering pursuit of emerging technology, high-quality engineers, processes, and practices. They help your business to leverage your data assets, creating the infrastructure, data culture, and technology ecosystems. Their Forecasting eBook teaches you how the Guided Analytics application combines just the right amount of automation and interaction for a specific set of problems. Learn more at www.knoldus.com

Open for Innovation KNIME

KNIME AG · Hardturmstrasse 66 · 8005 Zurich, Switzerland · info@knime.com · www.knime.com
©2021 KNIME AG. All rights reserved.
The KNIME® trademark and logo and OPEN FOR INNOVATION® trademark are used by KNIME AG under license from KNIME GmbH, and are registered in the United States and the European Union.

Try it out for yourself!

This workflow is available on the KNIME Hub:
tinyurl.com/forecasting-platform