Tangent Works

Time Series Machine Learning for the Future, Today!

Easy to use Predictive and Prescriptive Analytics

Digital Transformation through Machine Learning

Machine learning (ML) is the scientific study of algorithms and statistical models that computer systems use to effectively perform a specific task without using explicit instructions, relying on patterns and inference instead. Machine learning algorithms build a mathematical model of sample data, known as "training data", in order to make predictions or decisions without being explicitly programmed to perform the task. Machine learning algorithms are **used in a wide variety of applications**, such as email filtering, detection of network intruders and computer vision, where it is infeasible to develop an algorithm of specific instructions for performing the task.

Time series provide a completely different challenge than other ML tasks. In order to successfully make forecasts based on time series data, it is important to understand where the differences come from and how they can be exploited to create a fully automated modelling engine. Tangent Works worked hard to empower **TIM** (**Tangent Information Modeler**) to solve and overcome the inherent difficulties of working with time series data.

Use Cases

Time series are everywhere. From your local bakery trying to estimate the amount of pastries that will be bought next Sunday, to the weatherman telling you when you'll finally be needing those sunglasses and your bank trying to find out which interest rates will play out in their advantage: everything around you evolves through time. In many different industries, companies try to exploit time series data to generate the much sought-after **competitive advantage**.

Retail

Banking & securities

Healthcare

Transportation

Pharmaceuticals and life Sciences

Utilities

Services

Education

Manufacturing & Natural resources

Government

Time Series data is everywhere

In **utilities**, it is critical for many companies to accurately forecast electricity load. Electricity load is a fundamental input to operations of transmission system operators (TSOs) and is important for industrial producers to balance their decisions on electricity procurement. Owners of the photovoltaic plants, electricity traders and system regulators and need accurate forecasts of production of the photovoltaic plants, for different time horizons and different granularities, to optimise their maintenance, trading and regulation strategies; and the same goes for wind production. These examples only scratch the surface of the widespread presence of time series data in utilities.

Forecasting sales of a product or service plays an important role in the lifecycle of almost every **retail** company. The estimation of future sales can drive plenty of management decisions, such as efficient inventory management, preventing or early detection of potential issues, price setting and marketing. Another common task is to forecast future demand. For example, without a qualitative forecast, it can be very difficult to assume the right amount of stock that should be available.

In **finance**, entire companies are in business based on their ability to foresee future stock prices. Moreover, many financial institutions use their time series data in deciding future interest rates, in attempting to prevent fraud, in deciding which loan requests to approve and in many other scenarios.

These are only a few of the many use cases that time series data entail. Take a closer look at the operations in your own company, and in all probability you too will come up with some additional use cases.

Time Series Data Machine Learning

Machine learning is a buzzword right now. But it's not just about the machine learning technology; it's about the value it brings. Many of its concepts were developed as early as the 1940s and 50s. After the AI winter of the 70s, backpropagation — a mechanism to efficiently train neural networks — gave rise to various new use cases of ML. And now new insights in Machine Learning open new possibilities.

Some of these new possibilities are found in the field of **time series analysis**. Time series data is a series of data organised in time order. An important first form of analysis that is much requested relating to time series data, is **forecasting**. Apart from forecasting, many applications can be found for **anomaly detection**. Potential business use cases for time series data forecasting and anomaly detection are endless.

Model building

To benefit from **predictive and prescriptive analytics**, one needs to build models. This typically requires data scientists, who are very hard to find. A major trend in the market is to **automate the model building process**, freeing up time for scarce data science resources to work on data preparation and create business value. However, automating the model building process does not automate the crucial step of **feature engineering**. Especially for time series, feature engineering is a difficult task.

Forecasting

When forecasting future values for a time series, many aspects need to be considered, such as seasonality and dependency on feature variables either now or in the past. This creates **the need for a model** that helps in making forecasts. The **quality of the forecast** will depend on the quality of the input data and the quality of the model.

Anomaly Detection
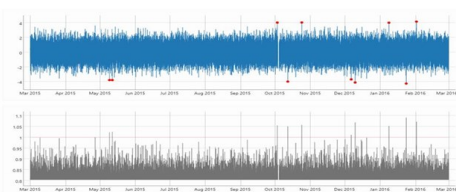
Anomaly detection entails **the identification of rare items or events**, i.e. data points with a suspicious behaviour in the data stream, including anomalies which are hard to detect by humans. This opens up a whole new range of applications in the field of **predictive and preventive analytics** including preventive maintenance, structural defect prevention and fraud detection.

TIM, Tangent Information Modeler

TIM, or Tangent Information Modeler, is Tangent Works' automatic model building engine. It is designed specifically for time series forecasting and anomaly detection. TIM has been tested numerous times and has proven to **generate high-quality models with zero degrees of freedom**, meaning no tuning of engine parameters is required from a user. Impressively, TIM in business user mode won the General Energy Forecasting Competition (GEFCom) in 2017, where it built over 150 models.

Tangent Works

Designed for Time Series

TIM is specifically designed to tackle **time series data**. Therefore, extra care has been taken to assure TIM is resistant to the **many pitfalls** of time series modelling. Find out more about these pitfalls and how TIM succeeds in overcoming them.

Changing data patterns

Phenomena represented by time series are **dynamic**. Consequently, a working model is not guaranteed to stay up to date. Numerous examples in different industries illustrate this reality: changing portfolios of assets in finance, changing influential factors in trading, where changes are so frequent it is mandatory to rebuild models from scratch, changing portfolios of production assets in the utility sector... Previously valuable models suddenly become useless in these new situations. This forces users to **repeat the model building process**, though this is clearly not an optimal solution as it can take much of the user's valuable time and effort. These new situations often require the **identification of new significant features** rather than a different modelling technique or slightly adjusted hyperparameters.

TIM empowers users to adapt to new situations by allowing models to be rebuilt or recalibrated continuously. Whereas **recalibration** only adjusts the model's parameters and leaves the model's structure (features) intact, **model rebuilding** starts by identifying new features and then builds a completely new model. Through the identification of new features, the process of model rebuilding is made robust to change.

Data Availability

Data availability introduces another complexity, as the **data availability situation** can vary from time to time. Certain data that was used for building the model, might not be available when a forecast needs to be made. As a result of this unavailable data, the model cannot be used for forecasting.

Date	Load	Temperature	Cloudiness	Irradiation
1/1/2012	14873.3	-1.3	4.1	0.2
1/2/2012	14253.3	-1.4	5	0.2
1/3/2012	13510	-1.4	5	0.2
1/4/2012	13053.5	-1	5	0.2
1/5/2012	12759.5	-0.6	5.9	0.2
1/6/2012	12697.3	-0.6	5.9	0.2
1/7/2012	12715.6	-0.6	5.9	0.2
1/8/2012	12457.7	0.8	7.3	0.2
1/9/2012	12297.9	1	7.3	0.2

Date	Load	Temperature	Cloudiness	Irradiation
1/1/2012	14873.3	-1.3	4.1	0.2
1/2/2012	14253.3	-1.4	5	0.2
1/3/2012	13510	-1.4	5	0.2
1/4/2012	13053.5	-1	5	0.2
1/5/2012	12759.5	-0.6	5.9	0.2
1/6/2012	12697.3	-0.6	5.9	0.2
1/7/2012	12715.6	-0.6	5.9	0.2
1/8/2012	12457.7	0.8	7.3	0.2
1/9/2012	12297.9	1	7.3	0.2
1/10/2012	12782.9	?	6.4	?
1/11/2012	?	?	6	0.2

In TIM, a **data availability scheme** is attached to each model building effort. This allows TIM to take expected data availability conditions into account when building models. Even so, it is possible that data availability changes over time or that some **unexpected changes** in data availability occur. Thanks to TIM's **InstantML** capabilities, a model can instantly be rebuild from scratch using the constrains of this new situation.

Multi-point Forecast

Situations where only a single-point forecast is required are rare in industrial practice. **Multi-point forecasts** are far more common, as they are required in many industrial verticals. Multi-point forecasts have traditionally been addressed by **multi-output models** and **recurrent strategies**.

Multi-output model

Intuitively, building and optimising a multi-output model is harder than doing so for a single-output model, because model parameters need to be optimised against all the outputs simultaneously. These models thus tend to exhibit a higher complexity compared to single-output models. Sometimes, this can even result in a **contradictory optimisation problem**.

Recurrent strategy

Recurrent strategies, on the other hand, optimise a single-output model that is then recurrently propagated in time, i.e. the forecast for y(t+1) is reused for calculating y(t+2). Recurrent strategies are, however, **prone to fast divergence**. This renders them impractical for a widespread industrial adoption.

TIM

TIM addresses multi-point forecasts by creating **a set of single-point models**. In this way, every forecast point has its own individual model that considers the corresponding data availability constraints. This set of models is referred to as a **ModelZOO**. Out of this ModelZOO, TIM automatically dispatches the correct model for the calculation of each point of the multi-point forecast.
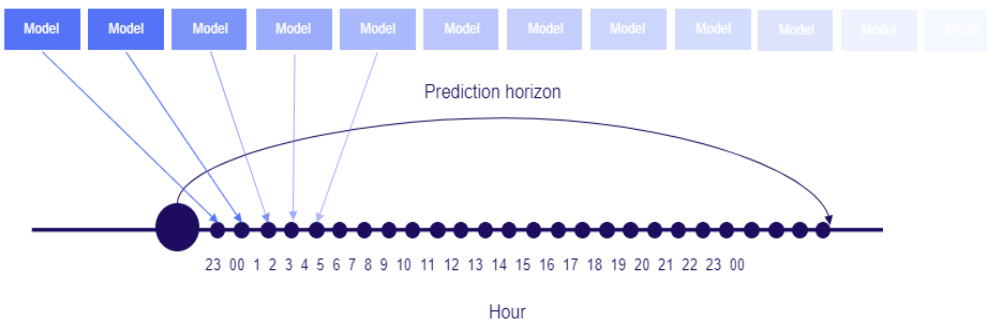
This approach has several advantages. First, by addressing each point individually, TIM can often achieve **greater accuracy**, because each model is optimised for a single point while considering the specific data availability. Secondly, each model's **features can be examined and compared**. As there might be rather different features driving different forecasts, this can provide **great insights**.

Multi-situational Multi-point Forecast

Often, multiple forecasts are needed at multiple points in time. In these cases, the user defines a **forecasting routine**, referring to a set of forecasting situations with their corresponding data availability schemes. In these cases, TIM's **multi-situational layer** assembles a ModelZOO that accounts for all the multi-point forecasts in all the situations of the user's forecasting routine.
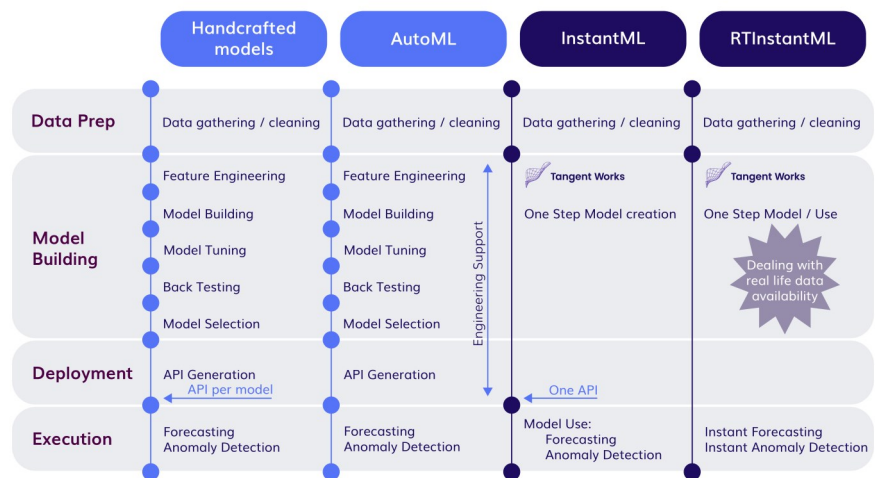
Tangent Works

- TIM (Tangent Information Modeler) Engine

- AutoML: time-consuming and expertise-intensive

- The importance of identification of significant features and overall modelling framework

- InstantML

- RTInstantML

- Anomaly detection

TIM, Tangent Information Modeler

Years of research have led to the creation of **TIM engine**, based on a unique blend of various advanced technologies including information criteria/ geometry—an interdisciplinary field that uses differential geometry techniques to study probability theory and statistics.

TIM differs to **Automatic Machine Learning** (**AutoML**) strategies that have recently drawn the attention of many researchers and industry professionals. AutoML focusses on the selection of an appropriate modelling technique and its hyperparameters for a task at hand. In this effort, AutoML solutions scan through many different ML libraries, create models and tune their corresponding hyperparameters. This has traditionally been a laborious task, remaining **time-consuming** and **expertise-intensive**.

Machine Learning Managing Complexity

	Handcrafted models	AutoML	InstantML	RTInstantML
Data Prep	Data gathering / cleaning	Data gathering / cleaning	Data gathering / cleaning	Data gathering / cleaning
Model Building	Feature Engineering	Feature Engineering	*Tangent Works*	*Tangent Works*
	Model Building	Model Building	One Step Model creation	One Step Model / Use
	Model Tuning	Model Tuning		Dealing with real life data availability
	Back Testing	Back Testing		
	Model Selection	Model Selection		
Deployment	API Generation / API per model	API Generation	One API	
Execution	Forecasting Anomaly Detection	Forecasting Anomaly Detection	Model Use: Forecasting Anomaly Detection	Instant Forecasting Instant Anomaly Detection

In time series modelling, we believe that **identification of significant features and the overall modeling framework** are far more important than the choice of a specific modelling technique and its associated hyperparameters. Applying this believe, we surpass AutoML with our unique **InstantML** technology for time series data. With InstantML TIM generates one single high-quality model with a single pass through the data. This modelling strategy identifies relevant features present in the data. And, it works significantly faster (seconds up to minutes on standard hardware) than typical AutoML strategies. The models are made available through **a single API**, thus saving organisations the hassle of creating separate APIs for each of the models.

Tangent Works has further enhanced TIM's capabilities with real-time InstantML (**RTInstantML**) technology that provides organisations with optimally trained models for their data available at any time with immediate forecasting or anomaly detection. TIM looks in the latest data and detects a corresponding data availability and forecasting routine out of it. There is no need to setup your data availability scheme for the situation – it is defined by the way data are organised and TIM will recognise this automatically. TIM will then **generate a model and produce a forecast in one single request**. The model is then discarded, as the next time a forecast is required the entire process is simply repeated. All TIM needs to know from a user is how many steps ahead the forecast should be calculated.

This way of working is especially advantageous if data availability scheme changes often or if forecasts are required ad-hoc, i.e. at any time of a day. RTInstantML does **not require any model management nor model storage**, which simplifies implementation complexity and shortens time to value. RTInstantML not only simplifies the model selection and automates feature engineering, but also tremendously shortens the lengthy process of developing the right model at the right moment, which is crucial for time series problems.

Automatic model generation for time-series goes hand in hand with **anomaly detection**. TIM generates high quality models over the in-sample training data and learns **normal behaviour**. TIM then checks for any significant deviations from this normality to find samples that are anomalous with respect to historical behaviour. This process is intuitive as TIM's models are **transparent** and describe the key drivers of the normal behavior. The degree of normality is provided as a quantity called the **TIM anomaly indicator**.

The Tangent Works Solution Architecture

The **TIM Solution** consists of three major parts:

The TIM Engine

The TIM Engine delivers all necessary **Machine Learning functionalities**, including time series model generation, forecasting and anomaly detection offering both InstantML and 'one call' RTinstantML with integrated model generation and forecasting. The TIM Engine can be used as a TIM **cloud service**, in your own cloud, **on premise** or on the **edge**. All of TIM's functionalities can be easily integrated through TIM's **open API**.

TIM Studio

TIM Studio is designed to serve as a **machine learning development and operations** (MLOps) **platform** for time series modelling. TIM Studio can be used by several types of people, ranging from business users, to citizen data scientists and even data scientists. It gives insight into the catalogue and versions of models, forecasts, detections… TIM Studio covers **data preparation**, **business process orchestration** including continuous data integration and much more.

TIM Clients

With the **'consume' strategy**, Tangent Works wants to make it easy for TIM to be used by anybody. To achieve this, TIM can be accessed through **a variety of tools and platforms**. TIM Clients are available not only for Alteryx but also for QLIK, Tableau, Denodo, Cloudera, and Microsoft or Microsoft Azure products, including Azure Data Lake, Azure Databricks, Azure SQL Data Warehouse, Microsoft SQL Server, Microsoft Excel, and Microsoft Power BI. This prevents users from having to switch tools each time a forecast or anomaly detection is required and seamlessly **integrates TIM's capabilities** into users' familiar digital environments.

Tangent Works

TIM and trends in the Data Science market

AI Service – In the Cloud

TIM cloud services are hosted services that allow applications, business productivity tools and platforms to incorporate the advantages of time series machine learning. Tangent Works can be deployed in the **public cloud**, making it extremely easy to use. If preferred, it can also be deployed on your **own cloud subscription**, **on premise** or on the **edge**.

Explainable AI

TIM generates **transparent models** that provide organisations with comprehensive **insights** into the models created along with helping them **measure the impact of predictors** on target values, automatically and instantly.

Edge Analytics

Edge computing is on the rise. With edge analytics, machine Learning modelling is brought closer to the sensors and **'Internet of Things'** devices. This offers opportunities to deliver **'swarm intelligence'** based solutions, offering collective behaviour based on distributed self-organized systems.

MLops

MLOps, **machine learning operationalisation** ensures versioning, release, activation, monitoring and overall governance of ML models. If you are only dealing with a few models, this is no big deal. When you really want to create value with machine learning, chances are that you will be using many different models. Tangent Works' TIM studio is a **setup studio** for business users and citizen data scientists and an **engineering work bench** for data scientists. Moreover, it provides **insight** into the models in catalogue.

Augmented ML

TIM delivers the value of machine learning without the complexity. TIM provides business users and citizen data scientists with a solution for making predictions, prescriptions and finding anomalies and outliers. TIM's augmented machine learning data science **automates fundamental aspects of advanced modelling**, such as feature and model selection.

Anomaly Detection

TIM delivers anomaly (or outlier) detection by detecting the **deviation from the norm**. The TIM engine uses the modelling capability to create a **normal behaviour** model. This is then **compared with the actual data** to detect early notice of potential anomalies. These functionalities allow for predictive maintenance scenarios, process monitoring, validation and much more.
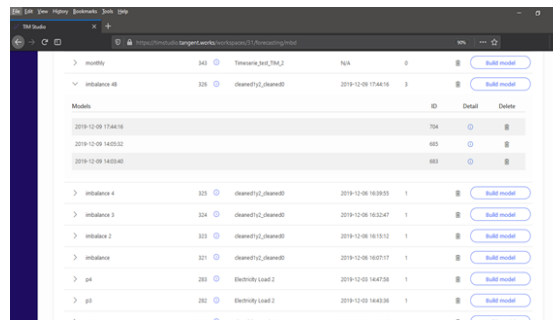
TIM Studio

MLOps and ML Workbench

TIM Studio is designed to serve as a **machine learning development and operations** (MLOps) **platform** for time series modelling. It is a **web interface tool** that allows users to manage datasets in workspaces, create and deploy models and examine results. This make is it a an excellent tool for both citizen data scientists and data scientists to create models, test results, start understanding their data...
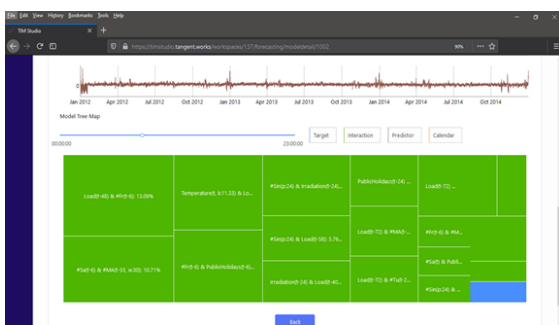
Model Catalogue

TIM shines in environments with many models, combining the strength of single-output models with the ability to generate fast and accurate models. TIM Studio gives users **a clear overview** of their datasets, models, forecasts and detections use both for experimenting and production purposes.


Explainable AI

To bring confidence to the use of machine learning it is vital to deliver **insights in the complexity and internal mechanics of the models** that are being generated. When analysing time series data the **feature engineering** is especially complex. TIM is designed to generate explainable models and TIM Studio builds on this to visualise what happens under the hood and offer **truly explainable AI**.

Forecasting

TIM Studio offers users easy to use functionality for examining input data and generating forecasts using **InstantML** and even **RTInstantML**. TIM Studio is also equipped with an **experimentation environment** allowing users to backtest their models and develop trust before deploying them in production.
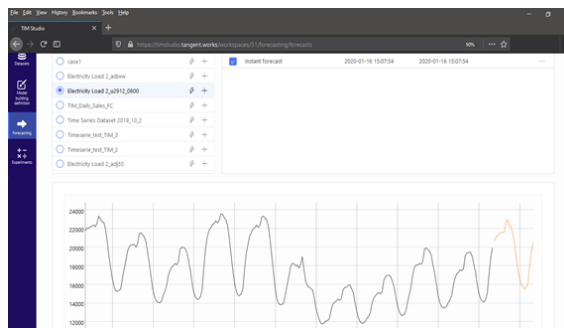
Anomaly Detection

Finding Anomalies becomes very straightforward using the detection functionality in TIM Studio. Users can choose to only fill in the desired **sensitivity** or to dive deeper into **the settings of the model building process**.

Tangent Works

TIM on Platforms

A set of **add-ins, tools, extensions and templates** is available to help you to start using TIM as quickly as possible, also on other platforms. Bring your data and let the TIM do the rest.

TIM in Excel

TIM in Excel is an add-in in Excel introducing **an intuitive interface to the TIM Engine in Excel**. The possibity to communicate with the TIM Engine from within Excel allows users to enjoy the capabilities and familiarity of Excel, combined with the unique time series insights realized with TIM. The TIM in Excel application supports TIM's unique technology called **RTInstantML**, allowing users to get direct forecasts based on the data in their Excel file.

TIM in Excel offers various benefits to **increase ease-of-use** of TIM's RTInstantML capabilities towards Excel users. For example, timestamps can be recognised in each of Excel's date-time formats, the dataset range is automatically extracted and additional comments or notes in the worksheet are automatically ignored. Moreover, users are free to choose any variable as target and they can select which of the predictor variables should be included in the forecast.

TIM in Alteryx

The **TIM Forecasting tool in Alteryx** offers an **easy to use, fast and accurate solution** for time series model generation and forecasting using **explainable AI**. Alteryx Designer empowers data analysts by combining data preparation, data blending and analytics; TIM expands this wide offer with advanced time series machine learning capabilities.

The TIM Forecasting tool in Alteryx supports TIM's unique technology called **RTInstantML** and also allows for the personalisation of the forecasting models through **adjustable advanced settings**. This tool also allows users to choose any variable as target and select which of the additional predictor variables should be included in the forecast.

As shown in the image below, the TIM Forecasting tool in Alteryx offers more **output channels** than solely the requested forecast. The output channels include the status of the request, the forecasted values, the importances, the extended importances and the graph. These channels integrate **explainability of the forecasting models** into the tool to truly give the user an interpretable view of what happens under the hood of the TIM Forecasting tool, namely automatic **feature engineering**, **feature selection** and **model building**. This explainability is further strengthened through a **visual representation** of these extended importances, allowing users to start understanding their data, the forecasting models and the actual
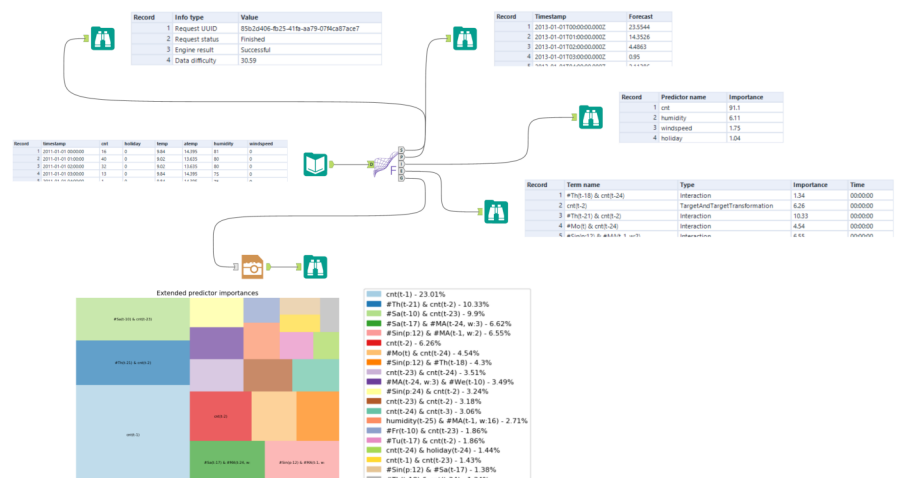
Data Integration Platforms

A prerequisite for leveraging machine learning is data. TIM integrates with a with range of **data integration tools** specialised in **gathering data**. The open nature of the TIM Engine makes it easy to integrate in any platform. This allows you to use TIM with your preferred environment.

Analytics and BI Platforms

Analytics and BI platforms deliver analytic content development by enabling nontechnical users to execute analytic workflows from data access, ingestion and preparation to interactive analysis. TIM introduces time series data **predictive and prescriptive analytics** capabilities in an easy, accurate, fast and explainable way.

IoT Platforms

The TIM technology can be embedded into **Internet of Things (IoT) devices or IoT Platforms**, equipping them with the **TIM Predictive and Prescriptive power**. TIM integrates with recognised IoT players.

Data Science Platforms

The distinction between BI and analytics platforms and **data science and ML platforms** is becoming less outspoken. Tangent Works' new way of dealing with **time series machine learning** is also available in data science ML platforms.

Cloud Platforms

The TIM Solution is a **containerised application** and can be deployed on any **cloud infrastructure**. Often, **Cloud Platform** vendors also have a rich set of data integration, orchestration and presentation functionalities.

Vertical Applications

Line of Business applications tailored to provide functionality for **vertical markets** can benefit from easy to implement time series machine learning implementing explainable AI without having to deal with the complexity and resource intensive machine learning expertise.

Tangent Works

Deployment

TIM can be deployed in various ways. The easiest way is to use the **Tangent Works Cloud**. TIM is **fully containerised** and can be setup to meet your **IT architecture requirements**.

Deployment Mechanism	When
SaaS Use the TangentWorks Cloud	• Easy • Safe • Scalable
Bring Your Own Cloud License Bring your own Cloud service to integrate TIM	• Have own cloud environment • Full infrastructure control
On Premise Install the TIM Engine container on your own Infrastructure	• Own Infrastructure, full control • Do not mind TCO aspects
On the Edge Install on an IoT device	• Bring ML to the data • Distributed architecture / Edge computing

About Tangent Works

Building and refining **accurate predictive models** was a laborious iterative task that required a combination of domain and data science expertise and weeks/months of effort. There just aren't enough data scientists in the world to meet the demands of a broad range of industries for which accurate and timely forecasts are essential to their competitiveness.

The founders' original concept was simple - what if we could build the model automatically by analysing historical time series data, including all the possible influencers, and figure out which data is relevant to the predicted output, in what combinations, and over which time intervals. From that **vision** and a breakthrough **architecture** based on a field of mathematics known as **'information geometry'**, the **Tangent Information Modeler**, **TIM**, was born.

TIM has since been recognised as **a winner in multiple competitions**, including the Global Energy Forecasting Competition, GEFCom 2017, and the 2017 ANDRITZ Hackathon, which was focused on optimisation and predictive maintenance for IoT applications.

Find out More

Documentation:
https://docs.tangent.works/

Try TIM on:
https://www.tangent.works/trial/

Tangent Works