WHITE PAPER

BayesCredit

Credit Risk Management and Credit Scoring

A functional and technical description of the approach and tool

COMBITECH HUGINEXPERT

WHITE PAPER

TABLE OF CONTENT

Contact information:

HUGIN EXPERT A/S: Anders L Madsen, Chief Executive Officer, +45 96 55 07 91

Combitech AS: David Häger, Chief Technical Officer, +47 988 800 50

Introduction

In spring 2015, HUGIN EXPERT A/S, the Danish provider of advanced decision analytic software, and Combitech AB, Nordic risk management advisory and consulting company, entered into a strategic partnership to strengthen risk management capabilities for clients who require advanced risk and security solutions. The partnership will help organizations accurately identify and manage risk, achieve compliance and avoid costly losses due to risk gone undetected.

Collaboration between Combitech AS Norway and HUGIN EXPERT A/S is not new. The two companies previously pooled their areas of expertise in a joint industry project to develop an operational risk management solution for a banking group in Norway that enabled bank compliance with BASEL II. Already then, the intersection of Combitech's deep industry knowledge and HUGIN's advanced decision support technology proved fruitful.

The objective of the latest strategic collaboration between Combitech and HUGIN is to combine their areas of expertise to help organizations meet their risk challenges and navigate in today's increasingly complex risk landscape, for example, organizations that must deal with money laundering, welfare risk, operational risk, credit risk and fraud detection.

A systematic credit risk solution

The key component of improved risk management is HUGIN's Bayesian Network modeling technology. The Bayesian Network model is a graphical representation of a client's entire risk scenario. The network model systematizes all available data and knowledge about a risk area - including conflicting information and unknown factors. After the model is constructed, the HUGIN software can make an accurate calculation of risk.

In this way the collaboration combines the deep industry insight of Combitech with HUGIN's sophisticated technology to create model-based solutions for clients that can handle even the most complex risk landscapes and are adaptable to the conditions and activities of the client organization – also as their needs change and evolve over time.

A credit-scoring model is the key component of modern credit risk management in which selected information is used to score a potential customer by calculation of probability of default (PD). A fundamental challenge of credit scoring models, and other models for risk assessment intended for high volume processing and assessment is the balance between information (including data) and credibility of results. It is not uncommon to design scoring models using large amounts of statistical material in the form of historical observations. Hence, the PD calculation for a specific customer is to some degree derived from the default statistics of similar customers, e.g., defined by address, income, social status and number of children. Furthermore, firsthand experience with the customer is also considered a valuable source of information. As a result, the initially assessed PD derived from expected values related to similar lenders becomes more specific to the individual in question when actual payment history becomes available.

As a result a bank may, for instance, be comfortable increasing the lending or credit limit for a customer that has demonstrated an excellent payment history and override the decision implied, e.g., by documented income alone.

The challenge of the credit scoring process is to use any available source of information to increase the sensitivity of the model so that it can identify good customers among customers that based on general data and averages would be considered poor. Enhancing credit model sensitivity with additional information increases the ability to compile a lending portfolio consisting of customers that represent acceptable credit risk as defined by the specific enterprise.

The flexibility in model design combined with capabilities to systematically and consequently introduce additional information to the analysis without corrupting the robustness of the data analysis is one of the key features of the HUGIN and Combitech credit risk management solution. The ability to assign a "correct" PD to a customer makes it possible for, e.g., a bank to take on customers that would normally not receive a loan due to conformance with a statistically defined category of poor lenders. Additionally, supported by a reliable and validated assessment of PD, the enterprise is provided with a strengthened ability to price credit more appropriately.

The ability to incorporate different sources of information into the credit scoring also provides the opportunity to generate decision support for a lending officer in a bank. For instance, in cases where a customer is assigned a high PD due to missing information or explicable deviations from the desired norm, the developed tool can assist the lending officer in identifying such rejections based on initial information and suggest additional information to confirm a decision to decline the loan or approve the loan. Hence, combining underwriting principles with limited manual processing strengthens the ability to identify good lending customers.

Proven solution

HUGIN software is being used by internationally active banks including the Danish Mortgage institution Nykredit. For almost two decades Nykredit and HUGIN have had a technology collaboration resulting in improved business efficiency in Nykredit's credit processes. In 2003 Nykredit launched BayesCredit – a solution based on HUGIN software which computes the probability that a corporate customer will default on its loan within the next year. BayesCredit proved to be extremely accurate and was accepted by Basel regulators as an advanced method for risk reporting. The solution was also tested by Standard and Poor's and found to perform better than Standard and Poor's own risk models when using the same input data.

The collaboration with Nykredit and other clients means that HUGIN consultants have a long-standing history helping customers implement efficient credit scoring models, and HUGIN software developers have experience in developing software solutions used for credit scoring.

Similarly, Combitech consultants have extensive experience with risk management including risk management related to credit processes. Several of Combitech's consultants have experience with developing risk models for assessing credit related losses using Bayesian Networks and HUGIN software. In addition, Combitech has experience from internal audit of credit processes including credit scoring, as well as performing research in the methodological foundation of credit scoring models. Using the HUGIN tool Combitech has also developed other risk scoring models similar in principle to credit scoring, e.g., customer classification models related to Anti-Money Laundering.

Combitech participated in a multi-client research project during 2007-2015 focusing on operational risk management and analysis in the banking and finance industry. Banks representing 80% of Norwegian banking capital participated. Credit risk is the main source of income for many Norwegian banks, making credit processes critical to business performance. Hence, these processes were subject to considerable analysis and modelling efforts during the project.

HUGIN delivered its advanced predictive analytics software to the multi-client research project in 2012-2014 when the developed models where operationalized. As part of this process HUGIN and Combitech collaborated on the use of the technology, as well as on new developments of the HUGIN software tool to meet the requirements of the end-user.

Together Combitech and HUGIN have a unique combination of technology, software, skills, knowledge and experience in using advanced statistical models to manage uncertainty in credit risk in an efficient and effective way.

Development process

The proposed system is an effective solution for default prediction based on advanced statistical graphical models. This document describes the Combitech process for developing credit scoring or default prediction models using HUGIN software, and how to integrate models into existing IT platforms using an Application Programming Interface (API) to the HUGIN Decision Engine.

Combitech has extensive experience with the credit risk management process, developing and implementing solutions for risk management in credit processes for financial institutions and providers of gambling services, and in collaboration with HUGIN has developed the methods framework and tools described here. When established, the described solution provides institutions with a genuine risk based approach for default prediction.

The development process is adapted to the specific needs of the customer and credit process to ensure that a credit scoring solution is adapted to the idiosyncrasies of the specific enterprise and industry. Solution development and implementation is normally carried out in accordance with the following sequential process:

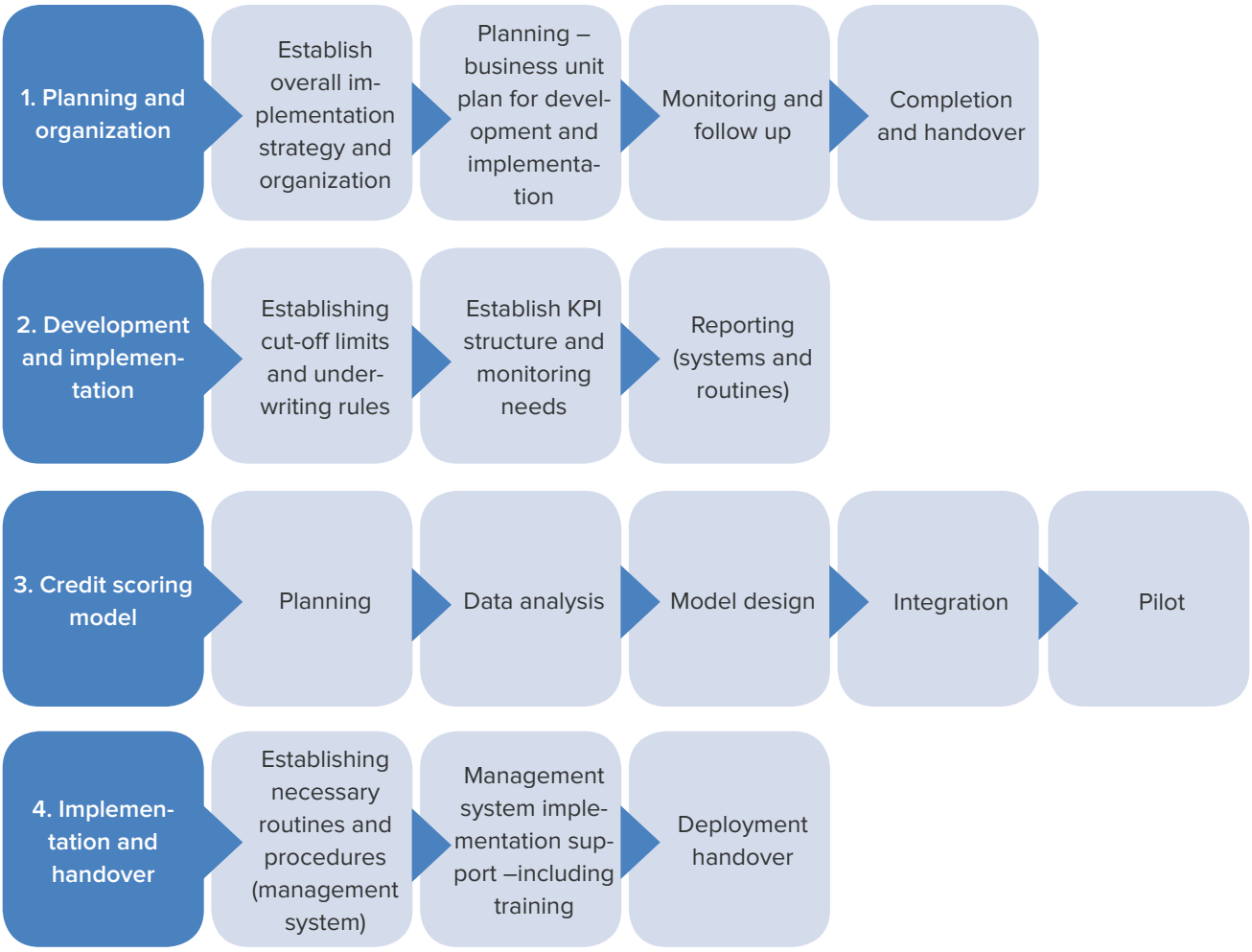


Figure 1: Step by step process for developing and implementing the solution

The planning of development and implementation activities includes establishing common rules and guidelines for development and implementation at the business unit level. The objective of the development and implementation activities is to ensure that models are developed in accordance with established specifications and requirements. Following a successful pilot, necessary governing documents (management system) are established, implemented (including necessary training) and the completed product/model is handed over to the client.

An overall project plan is established defining strategy and goals at a business unit level. This includes project organization, which needs to be synchronized with the enterprises' desired in-house control and ownership and the adherent internal organization and competence needs. The planning phase also includes establishing necessary communication routines, contacts and responsibilities.

Follow up and monitoring of project progress including KPI/KRIs should be established as well as a clear definition of each deliverable. Monitoring and follow up is established based on a project risk analysis in which identifying critical activities, deadlines and dependencies provides the foundation for assessing and managing events with a potential negative impact on project execution and goals. The objective is to ensure timely delivery of a system that conforms to performance and quality requirements.

The process of developing and validating the credit scoring model is performed in five steps including the activities accounted for in Figure 2. The process is also further detailed in the "Functional description" section.

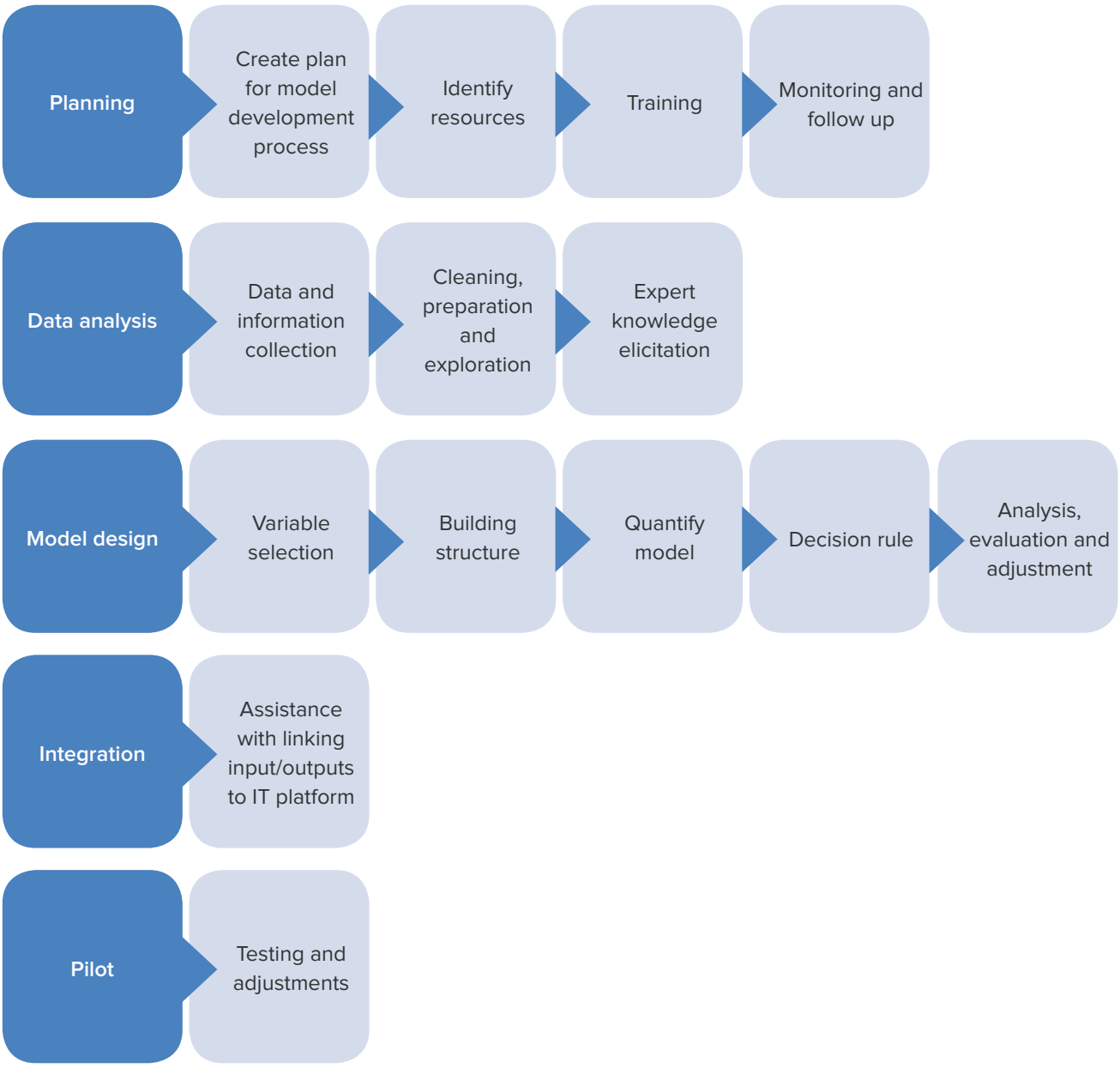
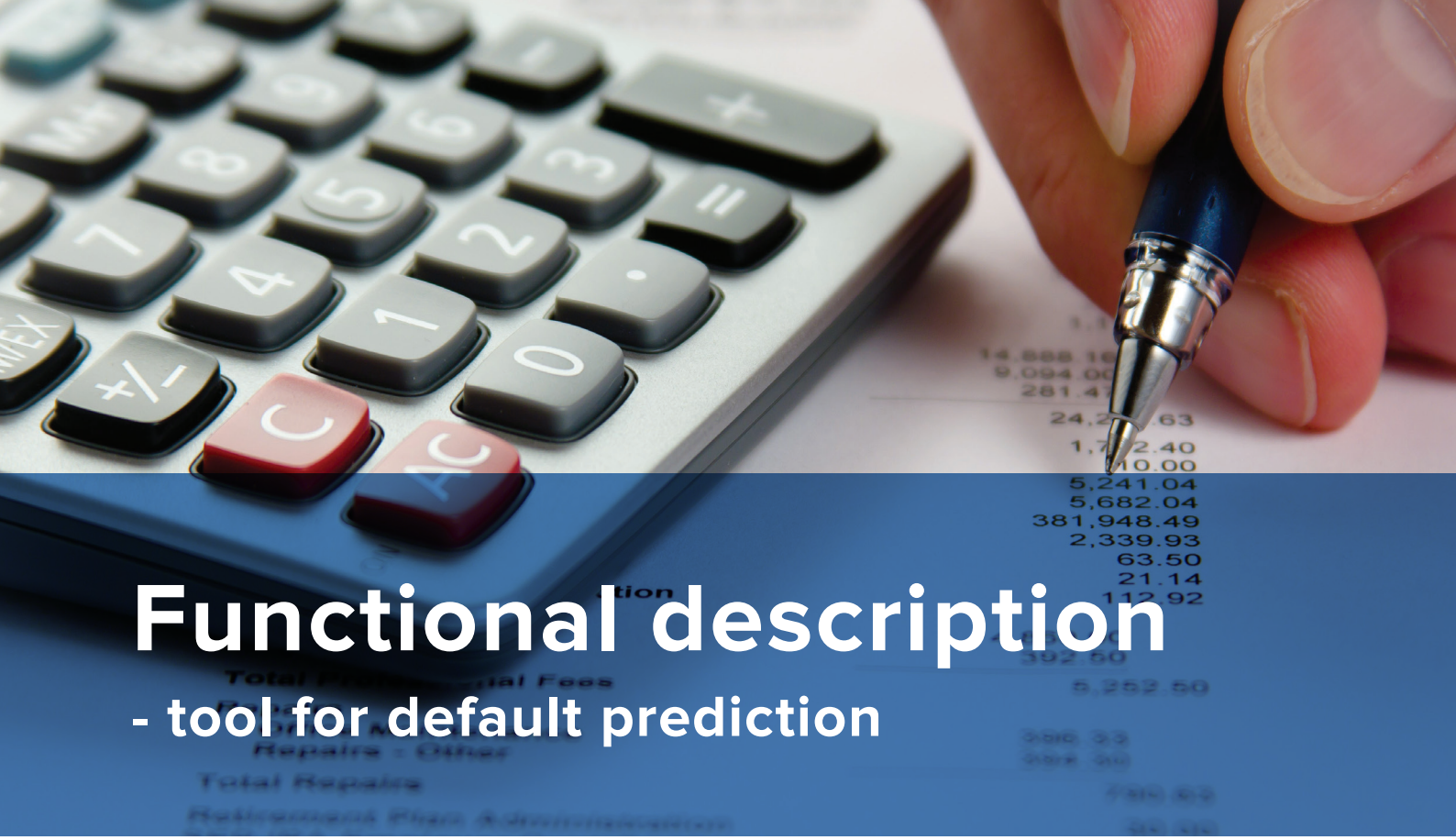


Figure 2: Step by step process for developing the credit scoring model

Once the plan for the model development task is complete, the model design itself starts with a review of available data. Here it is essential to account for the data available and to assess data quality. Furthermore, the data needs to be processed in order to establish a coherent dataset that can be used as model input. The analysis and processing of available data provides the foundation for assessing the need for additional information, e.g., in the form of expert knowledge.

Having established a thorough overview of available data, the model design process is initiated by the selection of variables. Here it is important that the selected variables are sufficient to enable the model to distinguish between customers representing high and low default risk. An erroneous or insufficient selection of variables may reduce the models ability to capture differences in default probability. Building a structure essentially means establishing dependence relationships between selected variables. Quantification of the model is carried out using both statistical data describing customers that constitute high and low credit risk as well as credit policy rules (e.g., income is more important to assessing default rate than address).

The next step is to evaluate the ability of the model(s) to distinguish between high and low default risk. This test is carried out based on data, such as the customer data of a selection of customers from the client organization. Test results are the starting point for model adjustment/calibration and rendering of results, which ensures compliance with any existing laws and regulations.
The integration of HUGIN model(s) into existing client systems can be performed using, e.g., the HUGIN Java API. This is normally carried out by the client with support from Combitech or HUGIN. The starting point for this work is an integration description with supporting examples.



Functional description
- tool for default prediction

The models are developed using Bayesian Network modeling tools in the HUGIN software.

The main component of the solution for default prediction is the graphical model (the network) itself, which represents the credit risk analysis. In this example, the model(s) calculate the probability of a default event based on given information about customers, customer relationships, products, services, geography, etc. And while in principle, model(s) may contain as many variables/indicators as desired and have any graphic structure, experience has shown that even a system consisting of a simple model with a limited number of indicators produces highly accurate predictions. To ensure the relevance of the tool and its associated models, the model development process is based on the historical data and information from the specific company and adapted to suit its particular conditions and business activity.

The development of default prediction models is first done graphically as illustrated in Figure 3. This involves systemizing all available information about the factors shown to have an influence on the probability of default. The visualization of identified causal relationships provides important insight and awareness about the potential causes of a default event. In addition to providing a visual overview that can be used for documentation and training, the network is built on a powerful mathematical platform that is the brain of an automated ongoing monitoring of customers' risk of default.

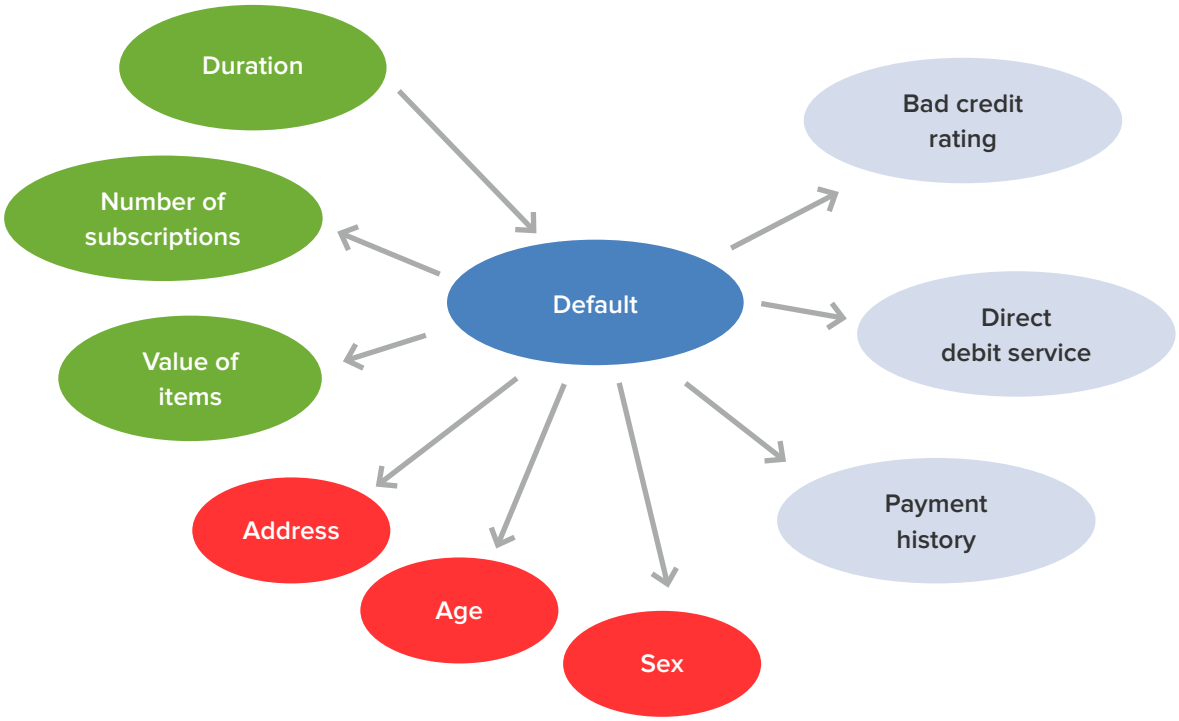


Figure 3: Graphical model of credit scoring

The system can be used to categorize customers into risk groups based on the established set of indicators. This risk analysis is the basis for credit granting decisions, and for follow-up and customer control measures to the actual risk exposed by the individual customer. Customers classified as extra high risk or high risk for default require enhanced customer due-diligence, whereas customers that fall into the normal risk category can be followed up with normal control measures. In certain cases, reduced risk may justify less rigorous controls and simplified due diligence measures.

When establishing a new business relationship or updating customer information, knowing the default risk supports the user, e.g., a sales person or lending officer in the credit process.

Figure 4 below illustrates a simple overview of a general credit process, including the decisions related to whether to approve or decline an application for credit. For new customers the scoring model provides a recommendation to approve or decline an application. The recommendation reflects the risk appetite and risk tolerance of the organization by declining customers that are considered too risky, and providing the appropriate risk class and implicit pricing for customers that qualify for approval. In addition to scoring the customer, there is also an opportunity to provide processing support to the employee handling the request for credit. For instance, the system can raise a flag when there is the potential for a false rejection in an application system (e.g., declined loans that could have been approved had the bank gathered further information), as well as indicate how the employee should efficiently confirm suspicions of a false rejection of the application. For existing customers an initial scoring is made using available payment history. Poor payment history may, according to established credit policy, disqualify a lender and lead to an automatic decline of an application.

On the other hand, excellent payment history may improve the credit score and justify an approval of an application that without the payment history would have been declined.

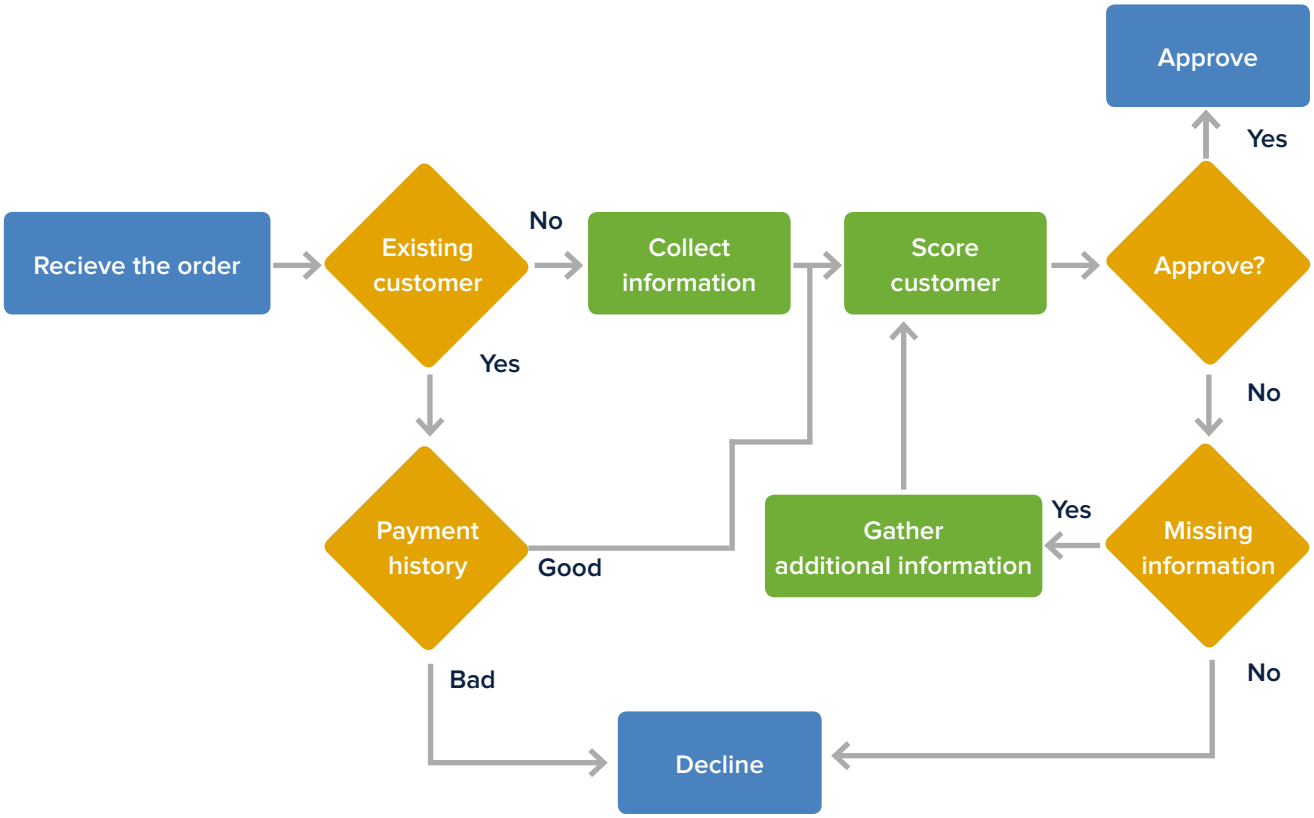


Figure 4: Description of functionality in the credit risk management solution

The customized default prediction models of the solution thus help lending officers or sales-staff manage customer relationships through the risk classification of customers, and provide guidance for establishing a level of pricing that corresponds to the risk indicated.

The below figure shows an example of a web-based interface of a credit scoring model:

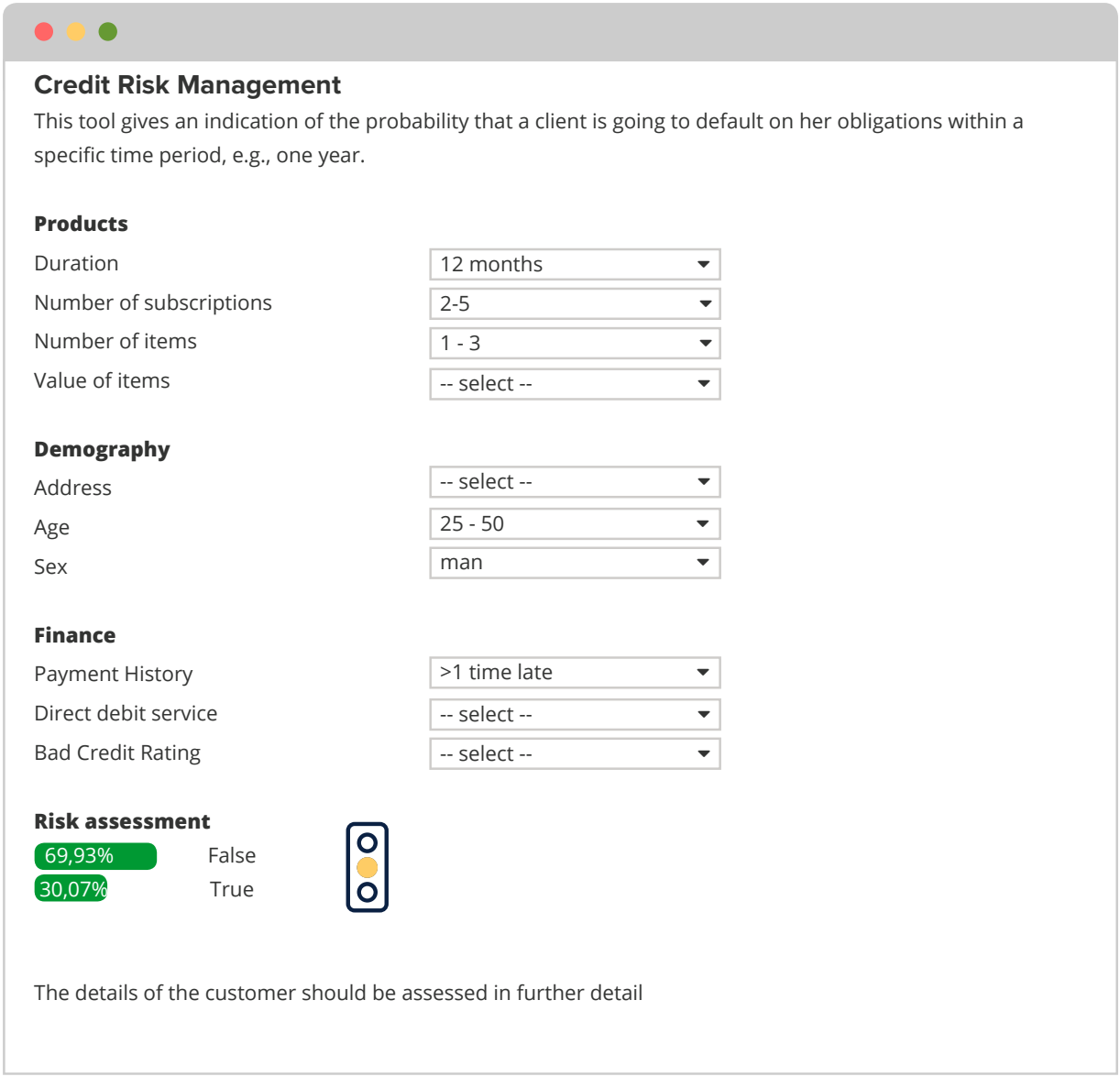


Figure 5: Web-based interface of the credit scoring.
See example at: http://demo.hugin.com/example/CreditScoring

The tool for monitoring customer and transaction default risk specifies the dependence relations between indicators of a default event and risk characteristics, and differentiates high risk customers from low risk customers. Further, the aggregate information and aggregate assessment of default risk based on a combination of customer risk characteristics, customer relationship, product and services and geography provide a solid decision-making basis for identifying high risk customers. It ensures resources are correctly dedicated to customers and transactions that pose the highest risk, and reduces the number of false positives and unnecessary use of resources.

Integration description

Here we describe how to integrate a credit scoring model into an existing IT platform. Integration involves two main tasks:

1. Linking each model variable to the appropriate data source.

2. Reading and presenting the probability of default as soon as new values are calculated.

Experience shows that integrating the credit scoring model into an existing IT platform is a simple task using a HUGIN API. Before integration it is important to decide how to display the probability of default to the user. We suggest showing the probability of default using traffic lights with red, yellow and green lights, or just red and green lights, as the threshold limits for when yellow turns to green, for example, are defined during the implementation phase.

Integration consists primarily of three steps:

1. Link each indicator variable to a data source in the existing IT platform.

2. Link threshold limit values for the state that indicates default to the traffic lights.

3. Implement a method and routine/process description that defines how to collect and register relevant data, update model probabilities and present results.

The below link contains a complete example of how to integrate a sample credit scoring model
http://download.hugin.com/Combitech/CreditScoring.zip

The example shown in the link uses the HUGIN Java API and involves the following steps:

1. Load the credit scoring model from file
2. Define how to obtain variables represented in the model
3. Compile the model into a calculation structure
4. Enter observations from a test example
5. Propagate evidence to calculate the probability of a default event
6. Present the results
7. Remove observations

Steps 1-3 are only required the first time. Steps 4-7 are performed each time the probability of default is calculated.

One of the benefits of the credit risk solution is the clear distinction between model development and model integration into an existing IT platform. Figure 6 illustrates the clear distinction between model development and system integration - usually performed by client IT staff and supported by HUGIN consultants if needed.

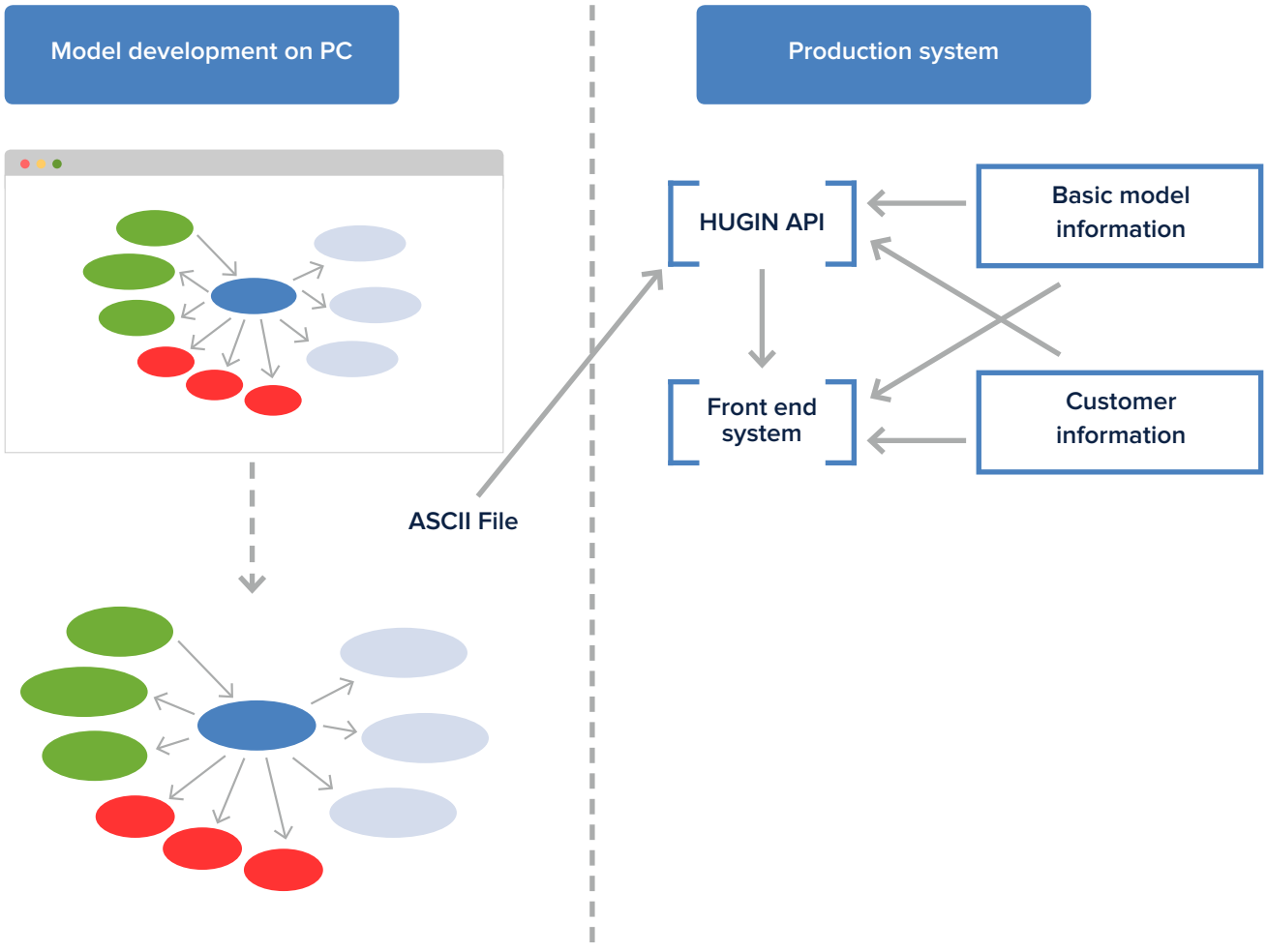


Figure 6: Clear distinction between model development and system integration.

After integration is completed, it is a simple task to replace an existing model with a new one. If no new variables are added to the model, the model can be updated by replacing a simple ASCII text file. If new variables are added to the model, these must be linked to the relevant data source.

As part of the development process, the HUGIN Decision Engine is integrated into the existing or new IT platform used by, e.g., a bank for managing credit. The HUGIN Decision Engine comes with a set of APIs which makes integration simple. However, it is common to build what is called a HUGIN driver that acts as a mediator between the existing IT platform and the HUGIN Decision Engine. The purpose of the HUGIN driver is to control the use of the HUGIN Decision Engine. The figure below illustrates this principle.

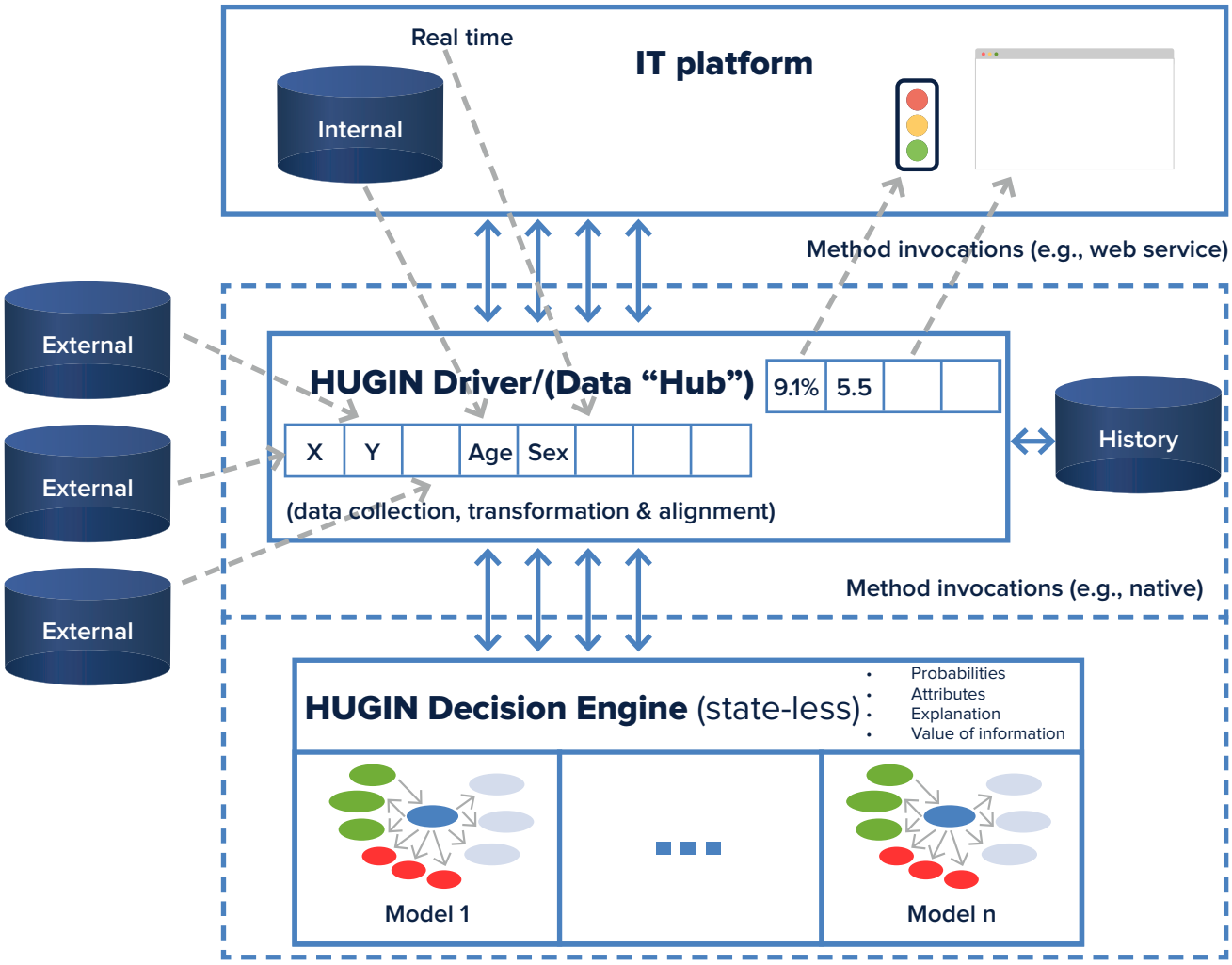


Figure 7: Illustration of Driver principle

The top part of the figure shows the platform's user interface for system users, and the bottom part the HUGIN Decision Engine. The layer in between is the HUGIN Driver, an application that can be developed by the client (or client's IT partner, assisted by HUGIN or Combitech if needed) to meet data format requirements. The task of the HUGIN Driver is to gather all data fed into the models and convert these into the proper data format (to match the data format in the models), and to convert the results produced by the HUGIN Decision Engine. The HUGIN Driver is also in charge of storing information in the form of data, which is registered in the models in addition to the model information used in each calculation. This historical data can be useful, for example, in testing future models.

See Chapter 1 of the HUGIN API Reference Manual (**http://download.hugin.com/webdocs/manuals/api-manual.pdf**) for further technical details about using APIs with other software and platforms, including information about how to compile and run applications using API functionality.

Technology

The technology behind the credit scoring models is Bayesian networks (also known as Bayesian Belief Networks, causal models of probability, Bayes nets, etc.). Bayesian networks are graphical models that structure the knowledge about a topic into a map of causes and effects between key variables. For each variable associated with a cause, probability is used to specify the extent to which one variable affects another.

HUGIN software has a long list of advantages compared to competing systems:

- It can combine expert knowledge and historical data
- It handles missing values and calculates using incomplete information
- Efficient inference engine makes real-time inference possible
- Models and implementation are flexible and easy to maintain, extend and revise
- The models are simple to integrate into existing systems
- Intuitive graphical models make it easy to communicate and discuss the credit scoring approach

As mentioned above both Combitech and HUGIN have significant experience in the use of the technology in credit processes.

Technical Information

Description of software

A HUGIN software package consists of the HUGIN Graphical User Interface (GUI) for developing models for credit scoring, and the HUGIN Decision Engine for integration. HUGIN Decision Engine has APIs for C, C #, C++, Java and VBA via a COM server. The HUGIN GUI is a Java implementation (which uses JNI) and the HUGIN Decision Engine is implemented using ISO C.

Description of data requirements

The HUGIN Graphical User Interface reads CSV files. The HUGIN Decision Engine has API functionality for entering data in real time. The HUGIN tool stores no data.

How the tool integrates with existing workflow

Credit models are integrated into existing systems using one of the API's for the HUGIN Decision Engine. The API's are available for C, C #, C++, Java and VBA through a COM server. The HUGIN Decision Engine contains functionality for entering information, making calculations and assessing the probability of default.

The credit scoring solution is especially useful in at least two different areas:

- For lending officers when registering new customers, during periodic review of existing customers and for training purposes. The credit scoring based on computing the probability of default is usually implemented as a traffic light/color indication on the user's screen.

- The graphical user interface is only used for developing models, during peer review, internal-external communication about the credit scoring solution, and training.

Technical requirements for required user interface

The HUGIN Graphical User Interface is implemented in Java using the HUGIN Java API for the HUGIN Decision Engine. The core of the HUGIN Decision Engine is implemented using the ISO C for portability and efficiency. HUGIN Decision Engine has been used in various hardware and operating systems, including PC s running on Windows and Linux, MAC for Mac OS, PCs and SUN servers on Solaris, and HP servers on HPUX and IBM mainframes.

Description of tool performance

HUGIN Decision Engine does not support user administration. Users of the HUGIN Decision Engine are responsible for using mutual exclusion variables, for example, when multiple users have access to the model at the same time. Alternatively, you could have a dynamic set of processes running on the same model. Time and space complexity for the most common credit scoring models are linear in the number of indicator variables. Response time is dependent on hardware.

Description of required administration of the tool

The HUGIN Graphical User Interface is installed on the desktop of PC users. The HUGIN Decision Engine is distributed as a single file or a few files installed on the host computer. There is no user administration. The credit scoring model is stored as a single file (usually as an ASCII file). This file can be opened in the HUGIN graphical interface and in the HUGIN Decision Engine. If minor adjustments are made to an existing model, the update consists of replacing the file on the server.

Management of software updates

The HUGIN graphical user interface (GUI) is a separate one-user program. The HUGIN GUI contains functionality for automatically installing new upgrades issued on the HUGIN website. A new version of the HUGIN Decision Engine can be installed by replacing the existing files on the host computer. Normally, new software versions are released 1-2 times per year, and active software users are notified of releases. The functionality of the HUGIN Decision Engine is tested using a series of test programs. Updates are not mandatory. New updates are usually installed on a host computer when new functionality is needed. HUGIN operates on the principle of reversed compatibility, meaning that files established in a previous software version will always function in new versions of the software.

COMBITECH

About Combitech AS

Combitech is one of the Nordic region's largest technology consulting firms, with more than 1,400 employees in Sweden, Norway and Finland. We combine technical expertise with in-depth industry knowledge, all-round capability and a particular focus on the environment and security. This results in customised, sustainable solutions for demanding clients who are active both in the Nordic region and internationally. Combitech is part of the SAAB Group with a global presence including 94 regional offices in 13 European countries. Our considerable presence in Scandinavia and Europe provides us with the necessary infrastructure to offer our services throughout Europe. For more information visit **www.combitech.no**

COMBITECH AS Prof. Olav Hanssensvei 7A NO-4021 Stavanger Norway
Phone: +47 475 11 275 www.combitech.no

HUGINEXPERT

About HUGIN EXPERT A/S

HUGIN EXPERT A/S is a provider of predictive analytic software for building model-based risk analysis and decision support solutions that can handle uncertainty. HUGIN EXPERT A/S was established in 1989 and is headquartered in Aalborg, Denmark. Clients use HUGIN tools to create intelligent decision support solutions for fraud detection, credit default prediction, operational risk management, medical diagnosis, health monitoring, risk analysis, data mining, troubleshooting, safety assessment, forensic identification and more. For more info visit **www.hugin.com**

HUGIN EXPERT A/S Gasvaerksvej 5 DK-9000 Aalborg Denmark
Phone +45 96 55 07 90 Fax +45 96 55 07 99 www.hugin.com info@hugin.com