IME i pakiet ExplainPrediction: zastosowania

Mateusz Staniak

18 III 2019 r.

Prof. dr. Marko Robnik Šikonja

Full Professor

T: +386 1 479 8241

E: marko.robniksikonja@fri.uni-lj.si

Office hours: Tuesday from 11am to 12pm (room R2.06, 2nd floor, right from the elevator).

Room: R2.06

LINKS

SICRIS

DESCRIPTION

My research interests include machine learning, data mining, intelligent data analysis, cognitive modeling, artificial intelligence, and their applications. In machine learning and data mining I am interested in ensemble learning, feature evaluation, probability prediction, cost-sensitive learning, feature subset selection, regression, natural language processing, and constructive induction (recent papers, free software). I am (co)author of approximately 50 publications. My papers have more than 2000 citations according to Scholar. More about me is on my personal web site. I tweet as @MarkoRobnikS.

Marko Robnik-Sikonja
@MarkoRobnikS
Ljubljana, Slovenia

Searcher and researcher. Interested in science, #machinelearning, #datamining, natural language processing, literature, purpose of life.
fri.uni-lj.si/en/employees/m…

117	240	126
TWEETY	OBSERWOWANI	OBSERWUJĄCY

TITLE	CITED BY	YEAR
Theoretical and empirical analysis of ReliefF and RReliefF M Robnik-Šikonja, I Kononenko Machine learning 53 (1-2), 23-69	1898	2003
Overcoming the myopia of inductive learning algorithms with RELIEFF I Kononenko, E Šimec, M Robnik-Šikonja Applied Intelligence 7 (1), 39-55	471	1997
An adaptation of Relief for attribute estimation in regression M Robnik-Šikonja, I Kononenko Machine Learning: Proceedings of the Fourteenth International Conference …	463	1997
Improving random forests M Robnik-Šikonja European conference on machine learning, 359-370	225	2004
Explaining classifications for individual instances M Robnik-Šikonja, I Kononenko IEEE Transactions on Knowledge and Data Engineering 20 (5), 589-600	131	2008

Marko Bohanec, Mirjana Kljajić Borštnar, Marko Robnik-Šikonja: Explaining machine learning models in sales predictions. Expert systems with applications, 71:416-428, 2017

Explaining machine learning models in sales predictions

Marko Bohanec[a,b,*], Mirjana Kljajić Borštnar[b], Marko Robnik-Šikonja[c]

[a] *Salvirt Ltd., Dunajska cesta 136, 1000 Ljubljana, Slovenia*
[b] *University of Maribor, Faculty of Organizational Sciences, Kidričeva cesta 55a, 4000 Kranj, Slovenia*
[c] *University of Ljubljana, Faculty of Computer and Information Science, Večna pot 113, 1000 Ljubljana, Slovenia*

Offerings

B2B Sales Excellence consulting

Data Science Workshops

The Predictive Index®

Benefits

- Simplified sales stages.
- Efficient dialog about progress of the opportunities
- Improved data points for informed business decisions
- Data Science applied to B2B Sales Management – www.SalesNT.com

- Build predictive models with your data
- Discover which factors drive a specific business outcome
- Enable data as a new voice in business meetings
- Learn R programming language for an elegant modeling

- Align talents with right jobs
- Understand which behaviors mark your top-performers
- Avoid guesswork when managing, coaching and much more…
- PI CA – State-of-the-art test of general cognitive abilities
- PI CA – It takes only 12 minutes, 50 questions
- Online assessments available in 70+ languages

Taksonomia wyjaśnień

1. *Expressive power*
2. *Translucency* (przezierność…) – jak bardzo wyjaśnienie wnika w strukturę modelu (dekompozycyjna/pedagogiczna/połącznie obu).
3. *Portability* – jak wiele metod można wyjaśnić.
4. *Quality* – wierność / zdolność przewidywania zachowania modelu / jednorodność (podobne zachowanie wyjaśnienie dla podobnych modeli) / *accuracy* / zrozumiałość.
5. *Algorithmic complexity*

Metody

- IME = wartości Shapleya
- Explain:

The straightforward characterization of the i-th input variable's importance for the prediction of instance x is the difference between the model's prediction for that instance and the model's prediction if the value of the i-th variable is not known: $p(y_k|x) - p_{S\setminus\{i\}}(y_k|x)$. If this difference is large, then the i-th variable is important. If it is small, then the variable is less important. The sign of the difference reveals whether the value contributes towards or against class value y_k.

$$\mathrm{WE}_i(k,x) = \log_2\left(\frac{p(y_k|x)}{1-p(y_k|x)}\right) - \log_2\left(\frac{p_{S\setminus\{i\}}(y_k|x)}{1-p_{S\setminus\{i\}}(y_k|x)}\right) \quad \text{[bits]}.$$

$$\varphi_i(k,x) = \frac{1}{a!}\sum_{\mathcal{O}\in\pi(a)} \left(\Delta(Pre^i(\mathcal{O})\cup\{i\})(k,x) - \Delta(Pre^i(\mathcal{O}))(k,x)\right) = \\ = \frac{1}{a!}\sum_{\mathcal{O}\in\pi(a)} \left(p_{Pre^i(\mathcal{O})\cup\{i\}}(y_k|x) - p_{Pre^i(\mathcal{O})}(y_k|x)\right), \tag{4}$$

$$\Delta(Q)(k,x) = p_Q(y_k|x) - p_\varnothing(y_k|x).$$

where $\pi(a)$ is the set of all permutations of a elements and $Pre^i(\mathcal{O})$ is the set of all input variables that precede the i-th variable in the permutation $\mathcal{O}\in\pi(a)$. This approach can be efficiently implemented and is called the IME method.

The method iteratively samples the space of attribute combinations (one iteration for each attribute set permutation $\mathcal{O}$). In each iteration, a part of attribute values is randomly generated (the attributes in $Pre^i(\mathcal{O})$) and the remaining attribute values are taken from the original instance x. The difference in prediction using either the original value or a randomly generated value for the i-th attribute is evidence for the importance of an i-th attribute in the interactions with other attributes. Details of this estimation procedure and its convergence can be found in (Štrumbelj and Kononenko, 2010).

O co chodzi?

Let us assume that the company is selling two complex solutions, A and B, on B2B markets. Their primary sales audience is business managers, and a certain level of sales complexity is expected (e.g., several business units are involved, the alignment of their expectations is not ideal, etc.). The company is successfully selling their initial Solution A, but recently Solution B was added to the sales portfolio. Ideally, the company cross sells the Solution B to existing clients. The sales personnel are not focusing their efforts on simple opportunities; rather they pursue deals in which they can offer complex solutions together with the company's deployment consultants. For a successful sale, the sales team attempts to engage senior managers at clients, with authority to secure funding and participate in the definition of requirements. Given its maturity, we expect sales of Solution A to be more successful than B. The company

Dane

Attribute	Description	Values
Product	Offered product.	Product code
Seller	Seller's name.	Seller's code
Authority	Authority level at a client side.	Low, Mid, High
Company size	Size of a company.	Big, Mid, Small
Competitors	Do we have competitors?	No, Yes, Unknown
Purchasing department	Is the purchasing department involved?	No, Yes, Unknown
Partnership	Selling in partnership?	No, Yes
Budget allocated	Did the client reserve the budget?	No, Yes, Unknown
Formal tender	Is a tendering procedure required?	No, Yes
RFI	Did we get Request for Information?	No, Yes
RFP	Did we get Request for Proposal?	No, Yes
Growth	Growth of a client?	Growth, Stable, etc.
Positive statements	Positive attitude expressed?	No, Yes, Neutral
Source	Source of the opportunity.	e.g. Referral, Web, etc.
Client	Type of a client.	New, Current, Past
Scope clarity	Implementation scope defined?	Clear, Few questions, etc.
Strategic deal	Does this deal have a strategic value?	Very important, etc.
Cross sale	Do we sell a different product to existing client?	No, Yes
Up sale	Increasing existing products?	No, Yes
Deal type	Type of a sale.	Consulting, Project, etc.
Needs defined	Is client clear in expressing the needs?	Info gathering, etc.
Attention to client	Attention to a client.	First deal, Normal, etc.
Status	An outcome of sales opportunity.	Lost, Won

Wynik modelowania

Table 3: The CA and AUC average performance on the business data set.

ML model	*CA*	*AUC*
RF	0.782	0.85
NB	0.777	0.83
DT	0.742	0.76
NN	0.702	0.70
SVM	0.567	0.59

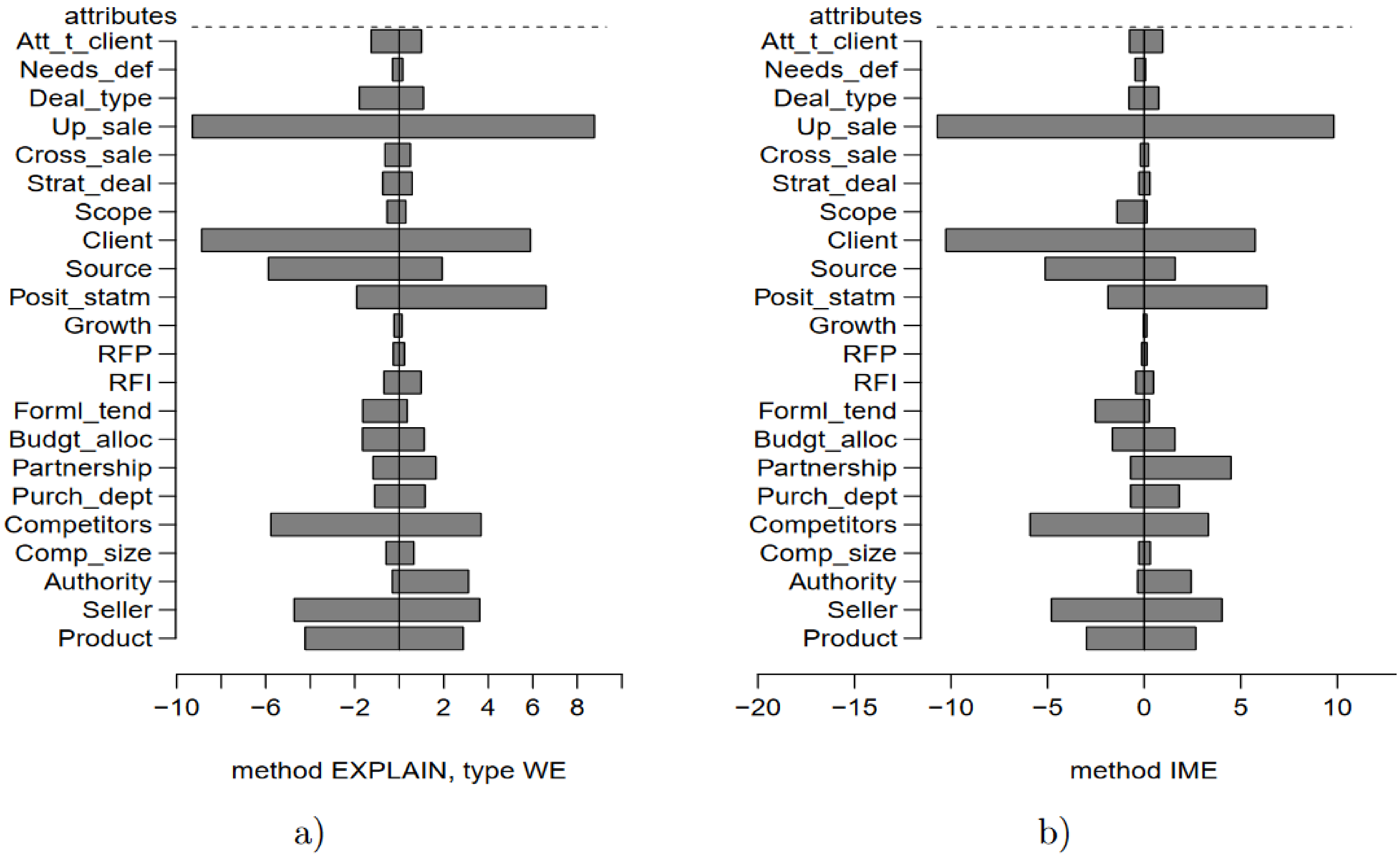


Fig. 4: Business case - EXPLAIN and IME model level explanations for RF model.

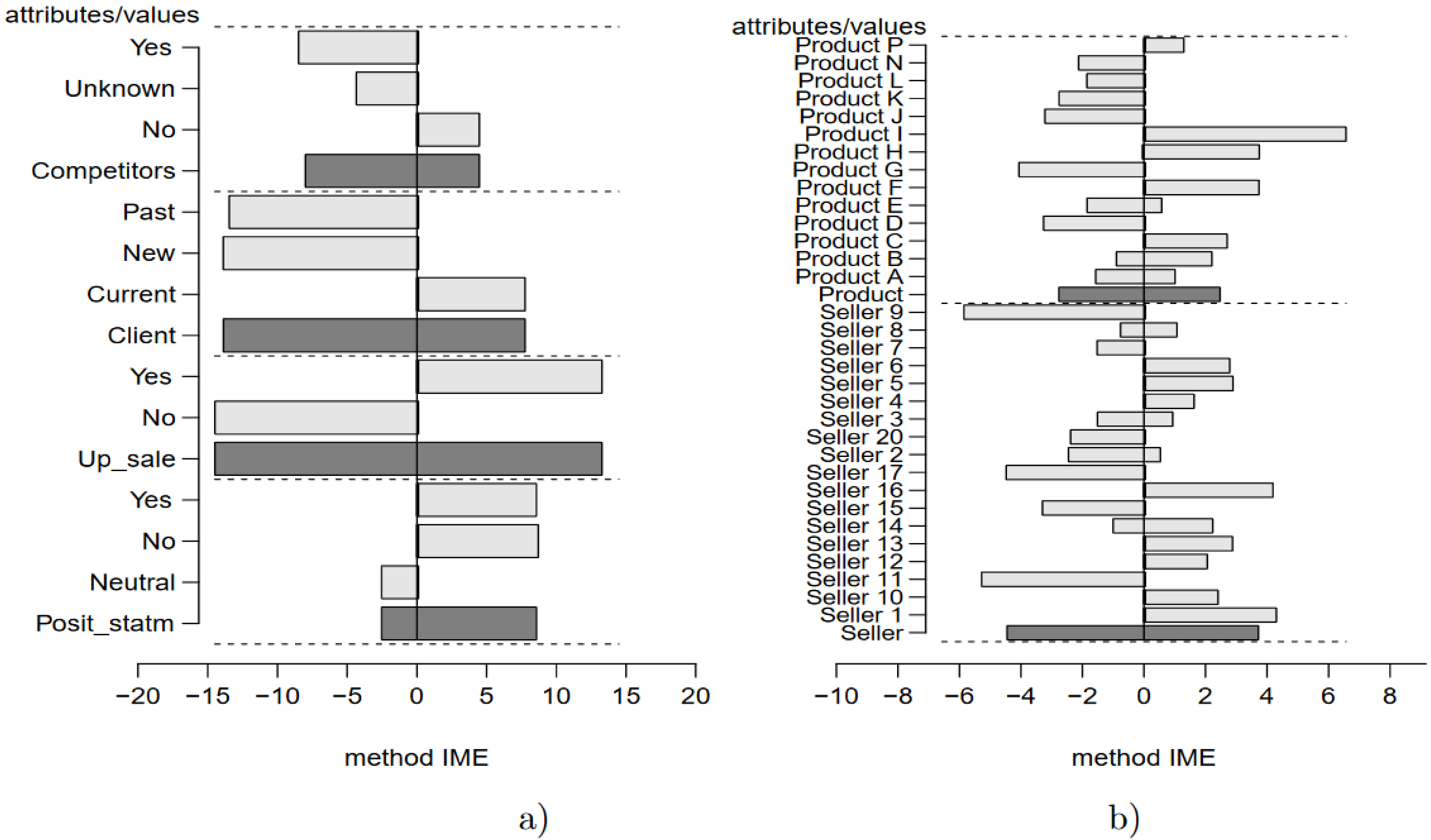
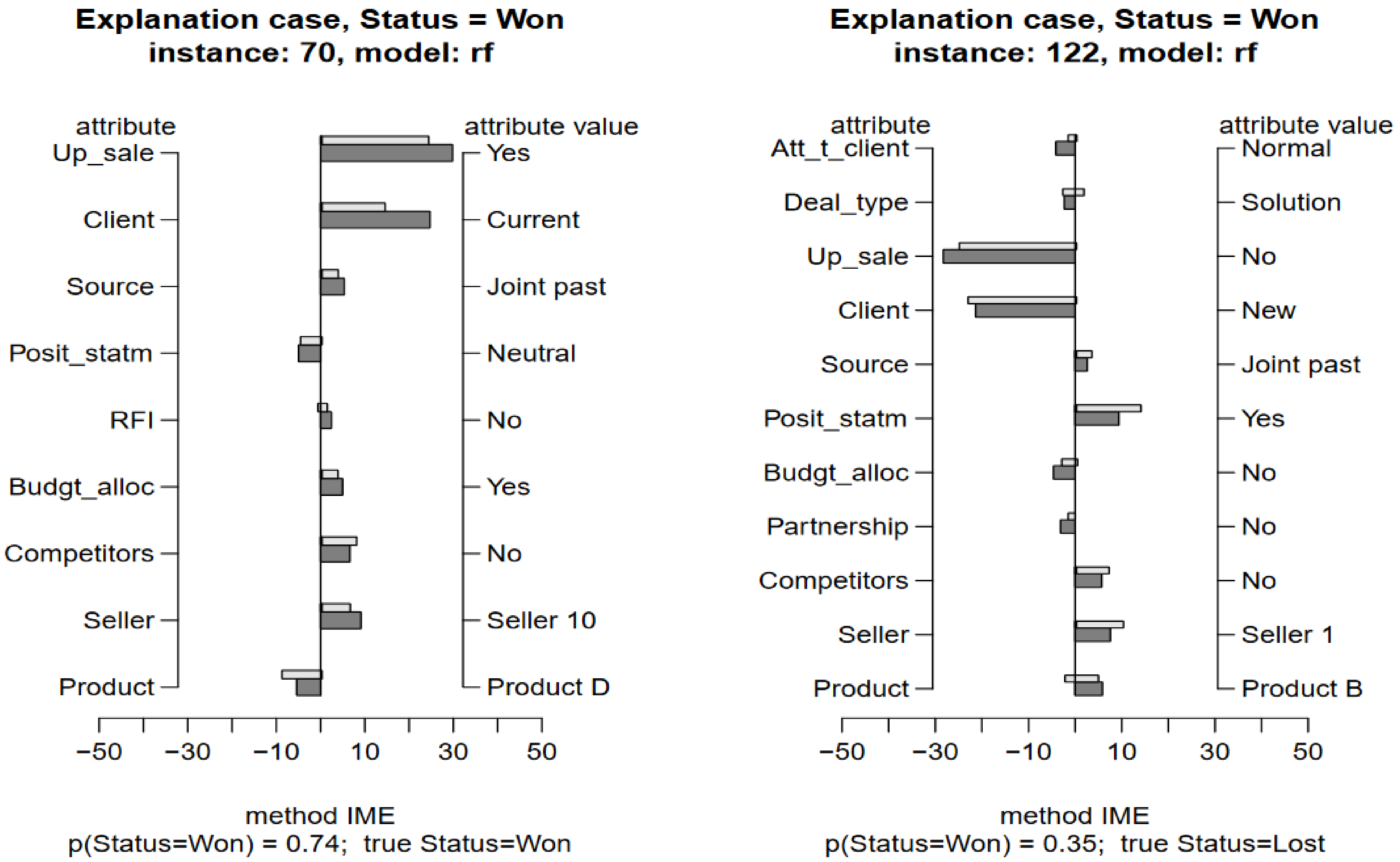
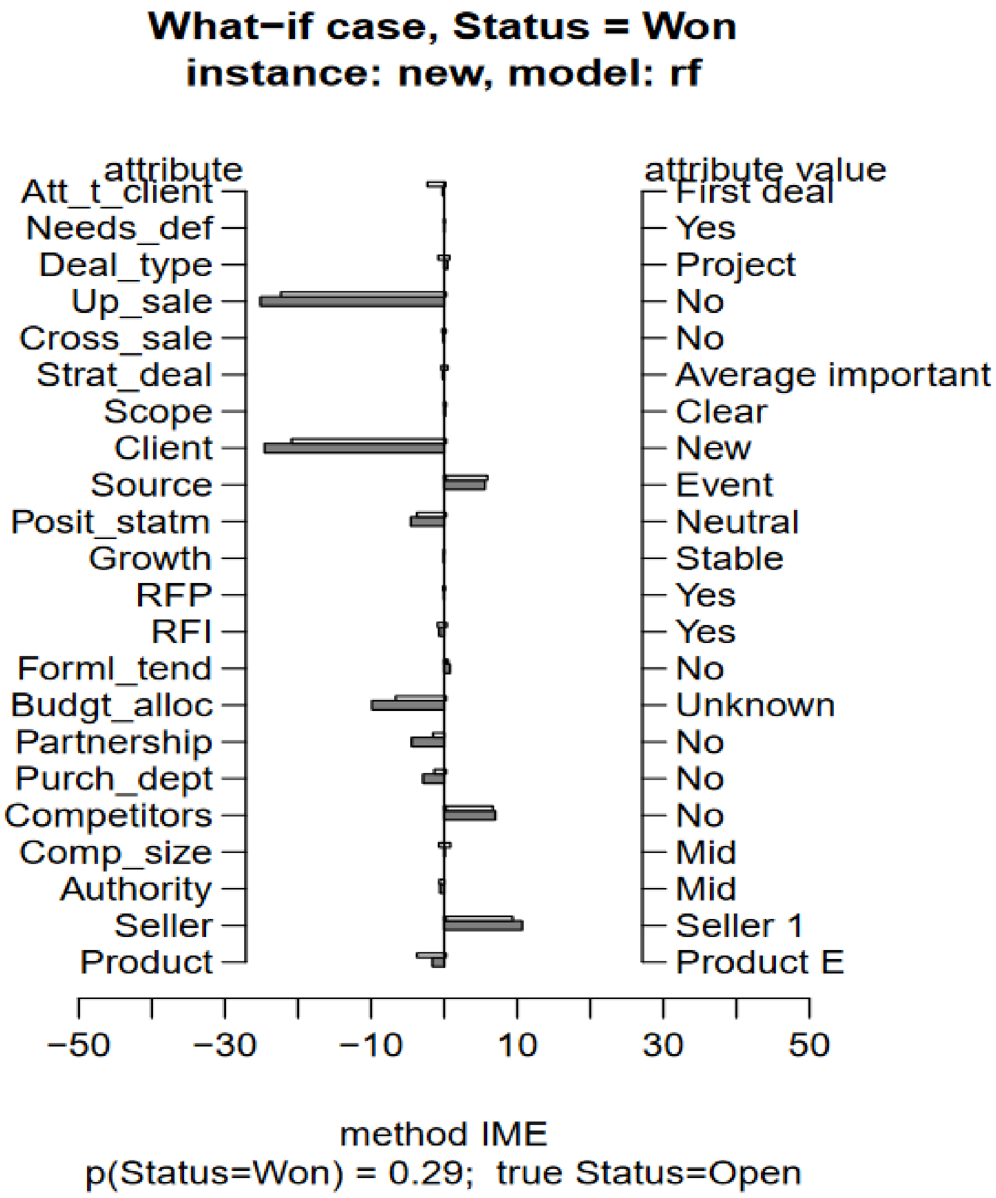


Fig. 5: Drilling into the model to visualize selected attributes and their values.

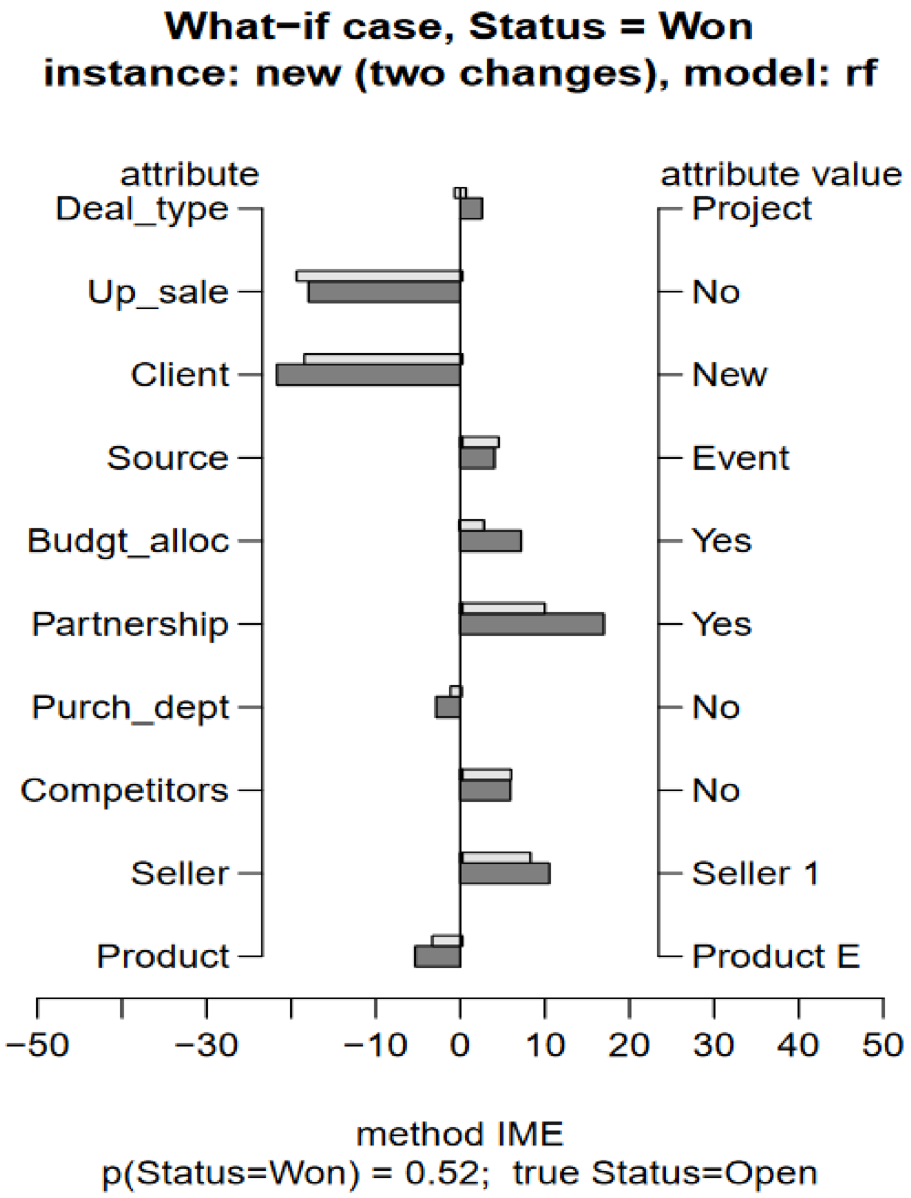


a) b)

Fig. 6: Instance explanations for one Won and one Lost deal.



a)



b)

Fig. 8: Initial explanation (a) and explanation after updates (b).

Wnioski

The presented intelligent system was tested for a longer period in a real-world company. The performance indicators confirm that forecasts created on the basis of the provided explanations outperform initial sales forecasts, which is in line with the intuition that explanations based on data better facilitate unbiased decision-making than the individual mental models of sellers. To foster the usage of the proposed methodology, the external consultant initially supported and trained the business users on how to apply the predictive ML models and the presented explanations. These activities have been proved to increase adoption speed;

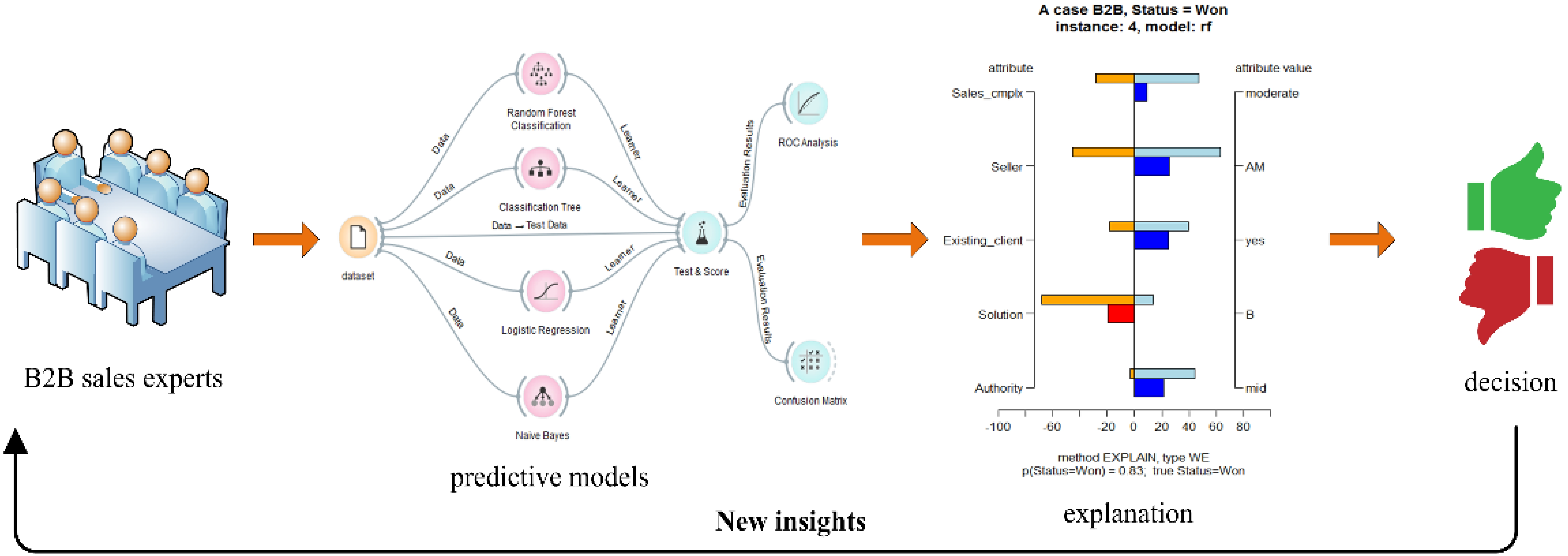


Fig. 1: High-level overview of presented intelligent system.

1. A group of sales experts is collecting historical B2B sales cases with known outcomes to support them in a prediction task for new sales opportunities. The collected data is processed by various ML techniques in the next step, resulting in the statistically validated prediction model. The explanation methodology provides explanations for the past and new cases and enables a cognitive evaluation of the model by the users. Based on the new insights, they have an opportunity to update the data set, retrain and re-evaluate the models before the next use, practicing human-in-the-loop in ML (Holzinger, 2016). As we are dealing with a

Human-in-the-loop (HITL) is a branch of artificial intelligence that leverages both human and machine intelligence to create machine learning models. In a traditional human-in-the-loop approach, people are involved in a virtuous circle where they train, tune, and test a particular algorithm. Generally, it works like this:

First, humans label data. This gives a model high quality (and high quantities of) training data. A machine learning algorithm learns to make decisions from this data.

Next, humans tune the model. This can happen in several different ways, but commonly, humans will score data to account for overfitting, to teach a classifier about edge cases, or new categories in the model's purview.

Lastly, people can test and validate a model by scoring its outputs, especially in places where an algorithm is unconfident about a judgment or overly confident about an incorrect decision.

Now, it's important to note that each of these actions are comprise a continuous feedback loop. Human-in-the-loop machine learning means taking each of these training, tuning, and testing tasks and feeding them back into the algorithm so it gets smarter, more confident, and more accurate. This can be especially effective when the model selects what it needs to learn next–known as active learning–and you send that data to human annotators for training.