Additional Experiments

Riccardo Galanti*‡, Bernat Coma-Puig†, Massimiliano de Leoni‡, Josep Carmona†, and Nicolò Navarin‡
*myInvenio, Reggio Emilia, Italy, ‡University of Padua, Padua, Italy, †Universitat Politècnica de Catalunya, Barcelona, Spain
Email: riccardo.galanti@my-invenio.com, {deleoni,nnavarin}@math.unipd.it
{bcoma,jcarmona}@cs.upc.edu

This section is intended to briefly show and explain additional experiments that have been made on several public datasets, in order to show the general validity of our framework. The characteristics of the events logs are presented in Table I, while the results obtained on those datasets are shown in Table II.

A. Bank Account Closure

We show here the application of our framework to the other two activities that are of interest for the bank, *Autorization Required* and *Pending Request for Acquittance of Heirs*.

In particular, in Figure 1 are shown the explanations related to the *Authorization Requested* prediction. Here, most of the times the LSTM is predicting that the activity will not be performed, and this is correct since the authorization to proceed further is usually requested to a director in the early stages of the process, rather than at the end. In particular, activities performed by the resource BOC or by resources with role Back-office are all performed after *Authorization Requested*, and this is of course also true when these activities have been performed in previous steps. Instead, when a customer does not want to change the bank but wants just to close one of its bank accounts among different ones he owns (labeled as *Closure_Reason=2 - Keep bank account. Same dip*), most of the times a director's authorization need to be requested; on the contrary, when a customer decides to change the bank and close definitively its bank account (labeled by *Closure_type=Client Recess* and by *Closure_Reason=1 - Client lost*), then a director's authorization is not needed.

B. BPIC 2012

This is the dataset from BPI 2012 challenge, and represents a real-life log of Dutch Financial Institute. Figure 3 shows the application for the remaining time prediction in the bpi12 dataset. The biggest value is represented by *event!=A_PREACCEPTED*, meaning that this activity is particularly important for the model. As a matter of fact, when this activity is performed (event=A_PREACCEPTED can be seen in the heatmap as well), the offer has an higher chance to be accepted by the customer, leading to a process that is longer compared to when the offer is declined (when event=A_DECLINED a smaller remaining time is expected). Looking at the history of *event!=A_PREACCEPTED*, the first two indexes (#0 and #-1) are mainly found at the beginning of the case, meaning that not until this activity is performed, there are good chances that the offer will be rejected (*A_DECLINED*), causing a smaller remaining time prediction; on the contrary, bigger indexes (#-2, #-3, #-4, #-5) mean that we are currently in the acceptance process (that is also indicated by the activities *O_SENT*, *O_CREATED*, and *O_SELECTED*) and a smaller remaining time is predicted. Finally, states of the work item belonging to the application seems to be not relevant for the prediction.

The financial institute aims also to identify in advance if a loan will be declined (represented by the activity *A_DECLINED*). The heatmap related to *A_DECLINED* prediction (6) shows as before that the activity *A_PREACCEPTED* being performed or not is the most influential. We already mentioned that when the offer is declined the process is shorter and not until A_PREACCEPTED is being performed (as can be seen in the first row of the heatmap), there are higher chances that the offer is going to be declined; on the contrary, when *A_PREACCEPTED* is performed or when it has been performed in the past (as can be seen in the long tail of the heatmap), then the offer is predicted to not be declined. In general when information for the application have been filled (labeled as *W_Completeren aanvraag*), *A_DECLINED* is predicted to not happen (shown in the boxes with #-1 or lower index), but when information have just been filled (as it can be seen in the box with the #0 index) there is still a little chance that the offer is going to be declined. Finally, when the amount of the requested loan is low (labeled as *Low value of AMOUNT_REQ* there is less probability that the loan will be declined by the institute (-2996) compared to when the requested amount is high (labeled as *High value of AMOUNT_REQ*.

C. BPIC 2012-W

Figure 7 shows the application for the remaining time prediction in the bpi12-W dataset, that is the dataset derived from bpi12 challenge, containing only the states of the work items belonging to the application. Here the largest and most important values are represented by the activity attribute; please note that here *Activity* and *concept:name* refer to the same thing (and this is also correctly understood by the model, which gives them the same importance), so for simplicity we will address them as activity. In particular, *W_Completeren aanvraag* and *W_Nabellen offertes* (labeled respectively as *concept:name=W_Completeren aanvraag* and

concept:name=W_Nabellen offertes) are always performed in the initial part of cases, and, hence, when they are currently performed (index #0) an higher remaining time is predicted. Instead, when these two activities have been performed in the past it means that either they are still performed because of internal inefficiencies that cause reworks or other activities are currently performed; in any case the process is not anymore in its early stages and hence a smaller remaining time is predicted (this can be noticed in the heatmap in the long tails of those two attributes). As a matter of fact, when these two activities are not currently performed (labeled as *concept:name!=W_Completeren aanvraag* and *concept:name!=W_Nabellen offertes*) it's even more influencial in predicting a smaller remaining time, because those inefficient reworks have already been done. According to the same logic, *W_Valideren aanvraag* and *W_Nabellen incomplete dossiers* are performed in the final part of cases and when it happen (labeled as *concept:name=W_Valideren aanvraag* and *concept:name=W_Nabellen incomplete dossiers*) a smaller remaining time is correctly predicted.

D. BPIC 2013

This is a dataset from BPI 2013 challenge, extracted from a company's incident management system. One question deals with the strategy of the company that most of the incidents need to be resolved by the first line support teams without involving 2nd or 3rd support line teams (push to front), leading to a more efficient process, so here we focused on predicting if at least one activity is going to be performed by a resource not belonging to the first line support team. The heatmap related to push to front prediction is shown in Figure 9; the fact that the country in which the problem is being addressed is Poland (labeled as *Country=pl*) is the most influential. The negative value -18295 means that the involvement of a 2nd or 3rd line resource is unlikely to happen. A further analysis of the data confirms this finding: when the country is Poland, then a 2nd / 3rd line resource is involved only in the 18% of cases. Values -4860 when the product is 424 (*Product=PROD424*) and -2528 when the product is 455 (*Product=PROD455*) indicate as well that the involvement of a resource not belonging to the first line support team is unlikely to happen. Conversely, when the country is India (labeled as *Country=in* with the value 10012), then an activity performed by a 2nd / 3rd line resource is more likely to happen; again an analysis of the data confirms that when the country is India a 2nd / 3rd line resource is involved only in the 79% of cases. Other important attributes positively contributing to predict an involvement of a resource not from the first line support team are related to the fact that the process is in the initial stages (labeled as *Low value of time from start* with the value 18093 in timestep 0 column) and that an activity is currently performed by a resource belonging to the 2nd or 3rd line support team (labeled as *Involved ST=2nd_3rd line* with the value 4504); note that the influence of these attributes decreases with the passing of time (see columns related to timesteps -1, -2, -3 and so on).

The company has also an interest in understanding if people working in the company is abusing the *Wait - User* substatus to hide inefficiencies in the process that would otherwise being detected by KPIs measuring the total resolution time of an incident. Figure 10 shows the results for the *Wait - User* substatus prediction. As before, if the process is in the initial stages the *Wait - User* substatus is of course more likely to happen, and this importance decreases with time; on the other side, if a *Wait - User* substatus is currently being performed (labeled as *Sub Status=Wait - User*) or has been performed in the past (see columns related to previous timesteps), then this activity is not predicted to be performed again. This substatus is similarly unlikely to happen if *Sub Status=Resolved* is currently performed, which is the activity that is set when the incident has been solved. Other important attributes are the country in which the problem has been addressed and the product that presented a problem or which had an incident; as an example, if the product is 158 (labeled as *Product=PROD158* with value 1650), 776 (labeled as *Product=PROD776* with value 1210) or the country is India (labeled as *Country=in* with value 1359), then a *Wait - User* activity is likely to happen. A further analysis of the data confirms this finding: when the product is 158, 778, or the country is India, a *Wait - User* has been used respectively in the 85%, 66%, and 63% of cases.

E. HelpDesk 2017

This dataset is a real-life log of SIAV s.p.a. company in Italy, representing instances of a ticketing process in the company helpdesk area. Figure 11 refers to the application for the remaining time prediction. Here, again, the fact that the process is in the early stages is the most influential factor. The positive value at timestep 0 (3129) indicates that the influence is towards increasing the value, namely towards having a largest remaining time. Other two important factors are related to the seriousness and priority of the ticket; if the seriousness is marked as 1 (labeled as *seriousness_2=Value 1* with value -2087), then a smaller remaining time is predicted, while if the seriousness is marked as 2 (labeled as *seriousness_2=Value 2* with value 1991), a larger remaining time is indicated. This fact is shown as well in the data, since when the seriousness of the ticket is marked as 1, the total time for the case to be executed is in general 6 days more compared to when the seriousness is marked as 2.

TABLE I: Event logs statistics

Event Log	# traces	# activities	mean events/trace	median events/trace	mean duration	median duration
Bank Account Closure	32429	15	5.5	7	15.5 days	7.7 days
BPIC 2012	12369	23	14	8	7.9 days	0.6 days
BPIC 2012 - W	9658	6	7.5	6	11.4 days	8.5 days
BPIC 2013	7554	13	8.7	6	12.1 days	7.6 days
HelpDesk 2017	4580	14	4.7	4	40.9 days	39.9 days

TABLE II: Event logs and summary statistics

	Remaining time MAE (days)	**Activity occurrence prediction (AUROC / APR / F1)**			**Cost prediction MAE (Euro)**
Bank Account Closure	4.37	Authorization Requested 1 / 0.99 / 0.99	BO Adjustment Requested 0.86 / 0.69 / 0.65	Pending Request for acquittance of heirs 0.99 / 0.87 / 0.90	0.95
BPIC 2012	6.66	A_ACCEPTED 0.92 / 0.66 / 0.60	A_CANCELLED 0.74 / 0.52 / 0.37	A_DECLINED 0.77 / 0.55 / 0.51	-
BPIC 2012 - W	7.84	-	-	-	-
BPIC 2013	11.82	Wait - User 0.50 / 0.43 / 0.45	2nd / 3rd line 0.90 / 0.90 / 0.81	-	-
HelpDesk 2017	5.96	-	-	-	-

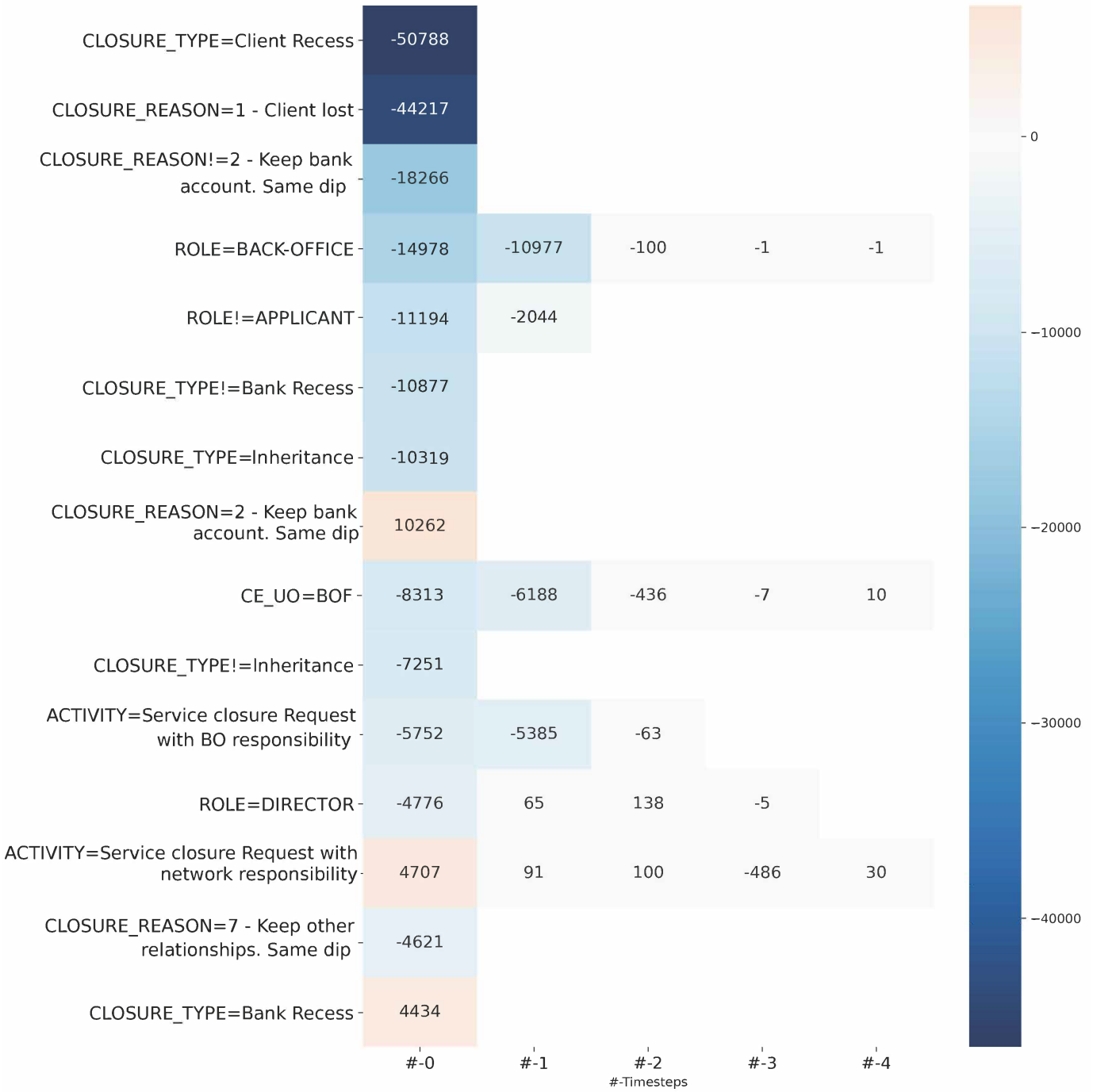

Fig. 1: Offline explanations for *Authorization Requested* prediction (Bank Account Closure)

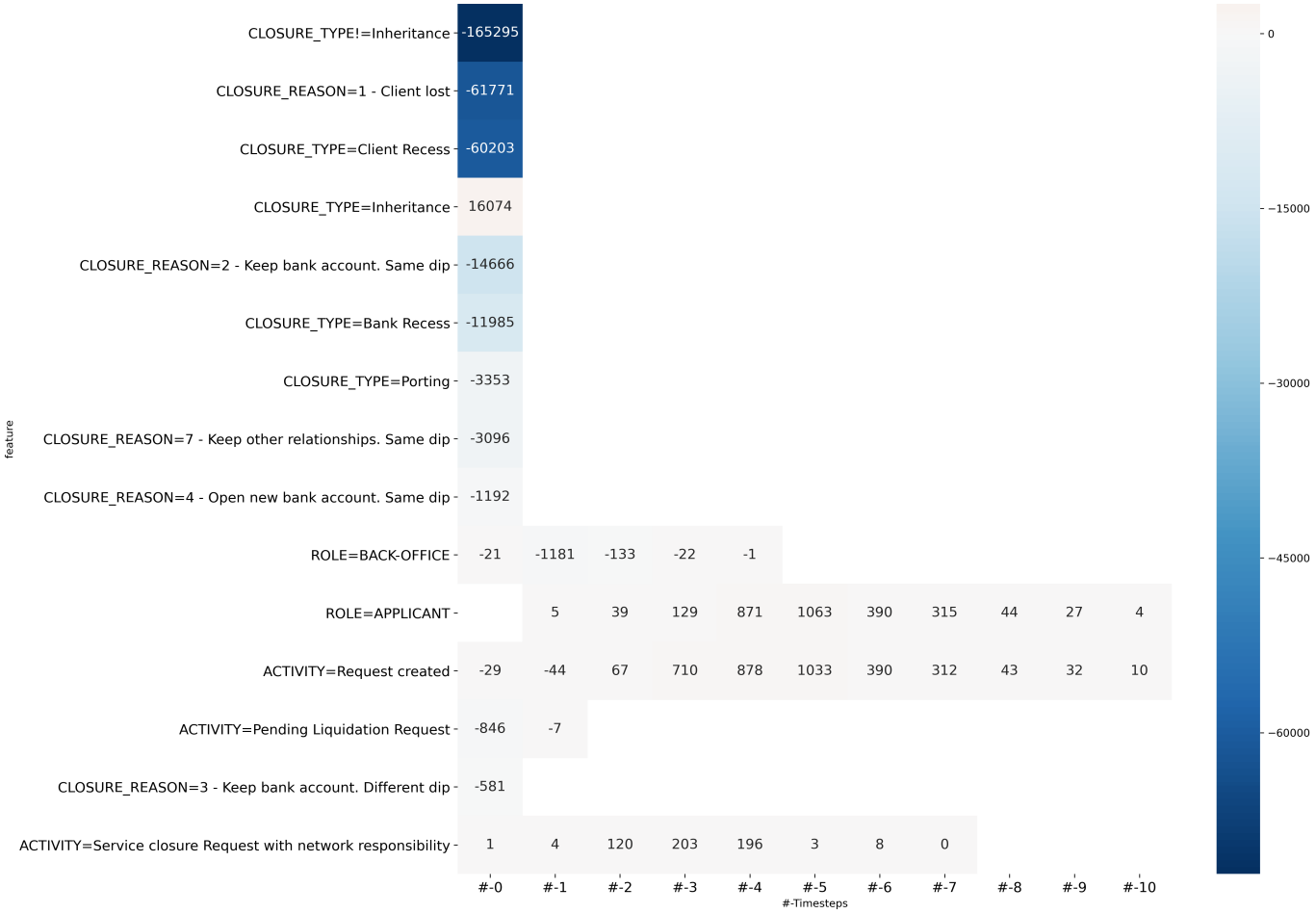

Fig. 2: Offline explanations for *Pending Request for acquittance of heirs* prediction (Bank Account Closure)

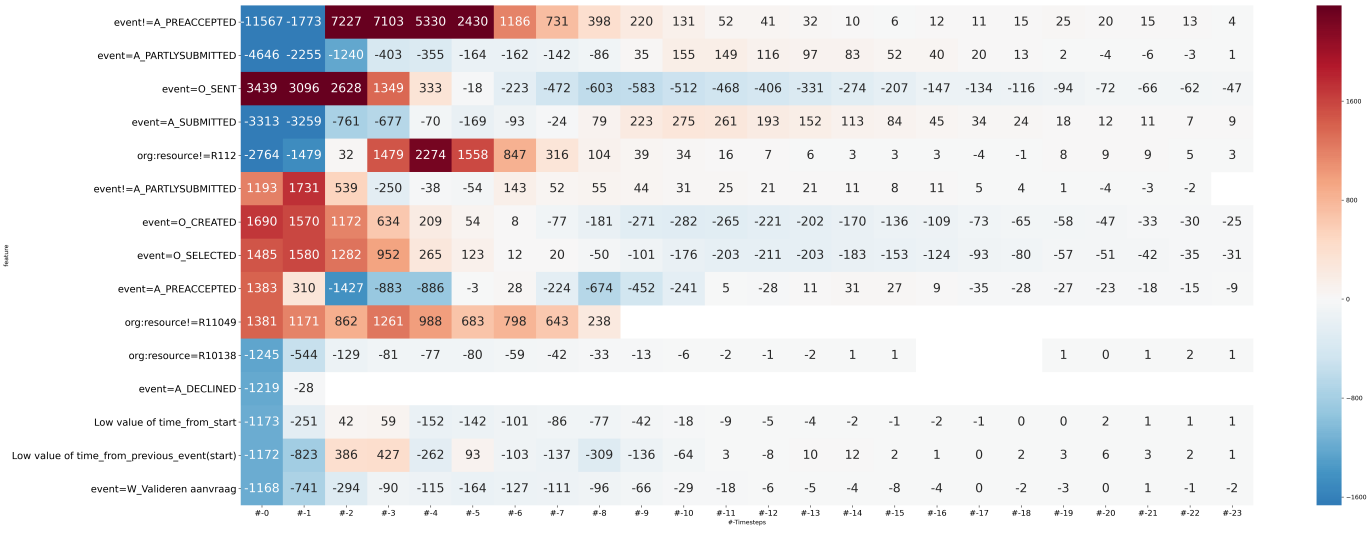

Fig. 3: Offline explanations for BPI12 remaining time prediction

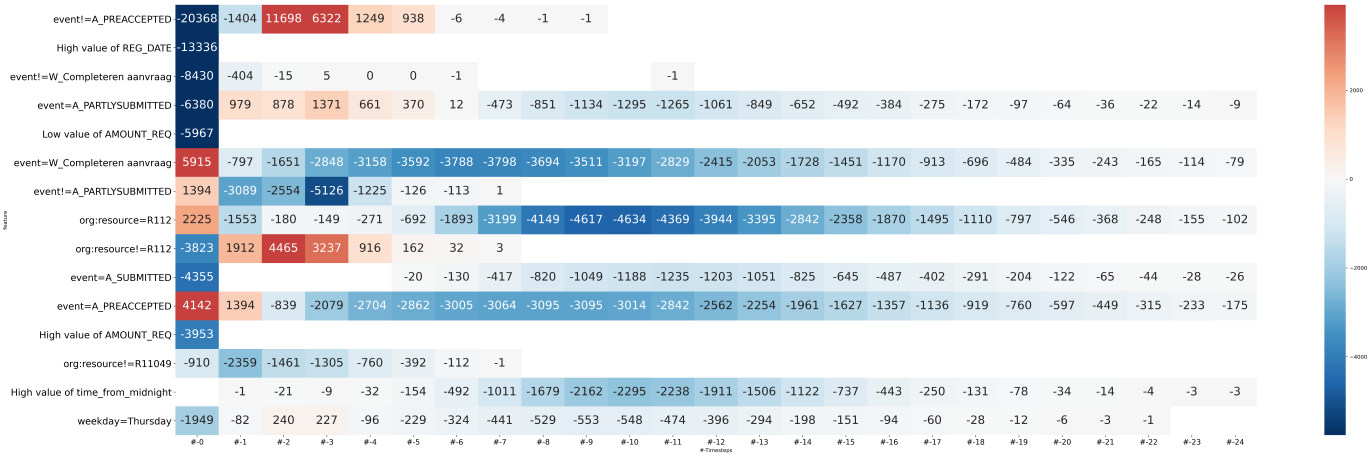

Fig. 4: Offline explanations for BPI12 *A_ACCEPTED* prediction

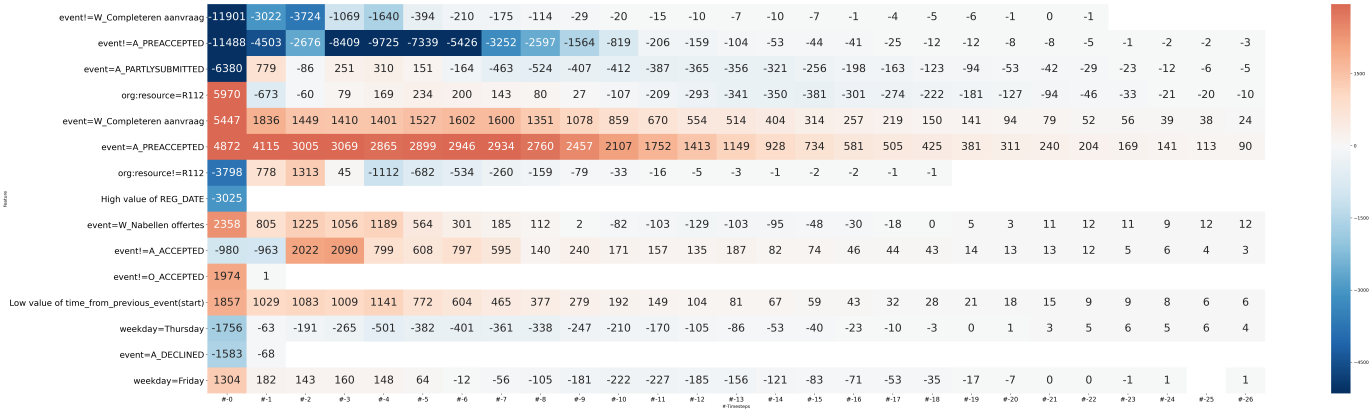

Fig. 5: Offline explanations for BPI12 *A_CANCELLED* prediction

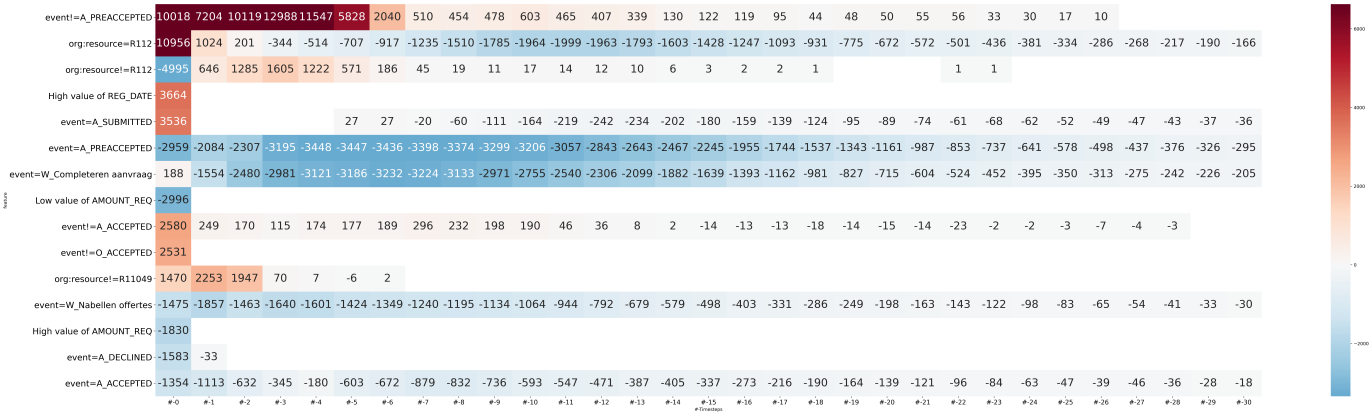

Fig. 6: Offline explanations for BPI12 *A_DECLINED* prediction

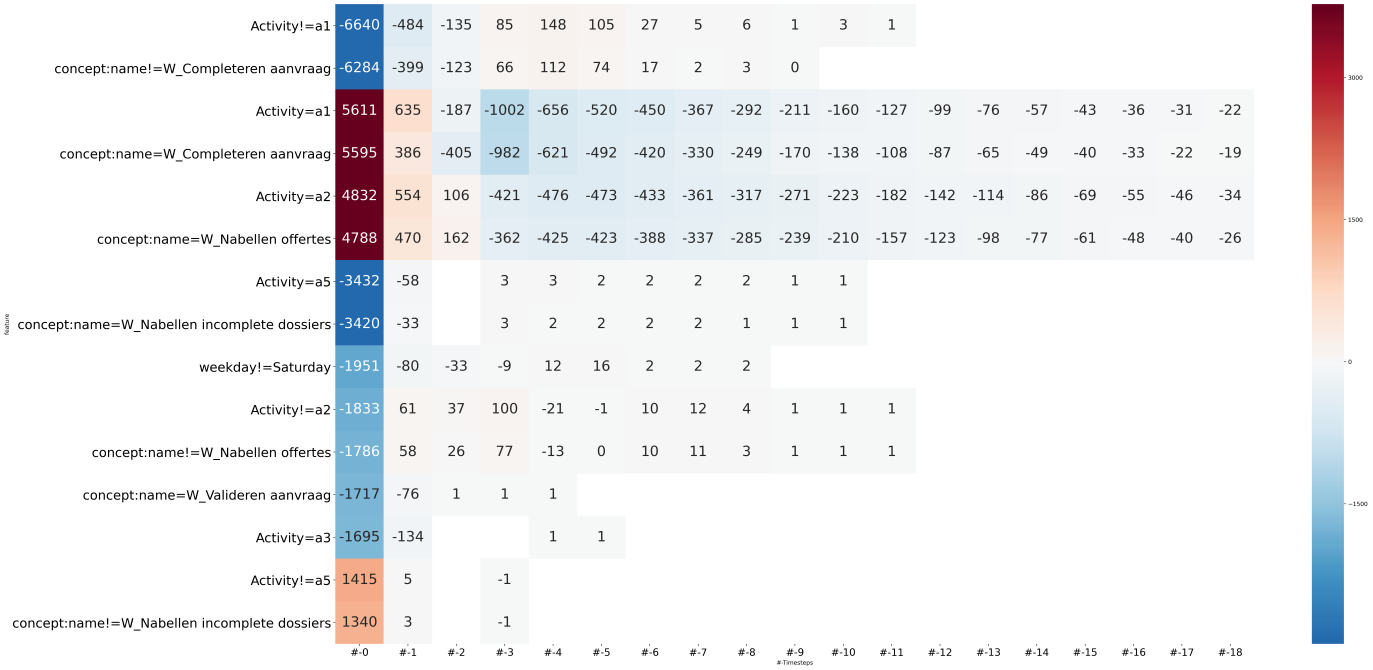

Fig. 7: Offline explanations for BPI12-W remaining time prediction

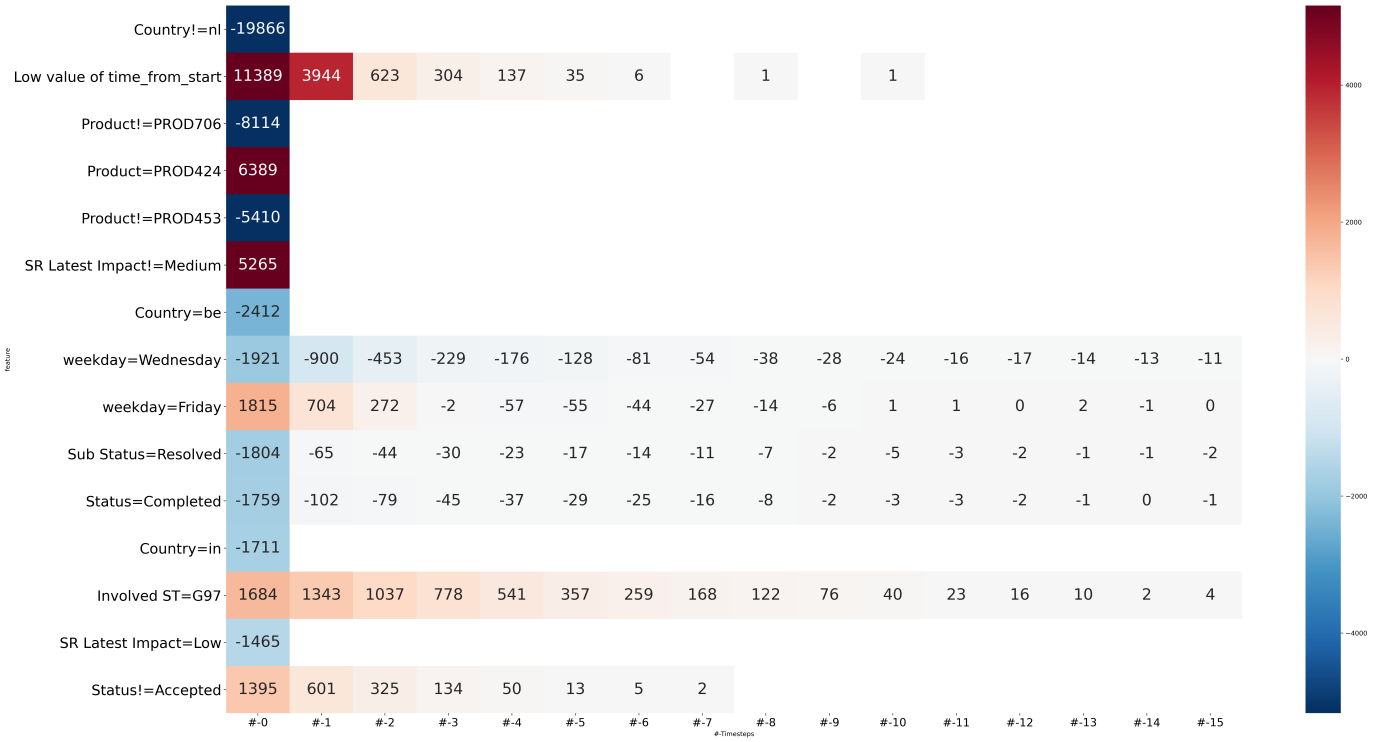

Fig. 8: Offline explanations for BPI13 remaining time prediction

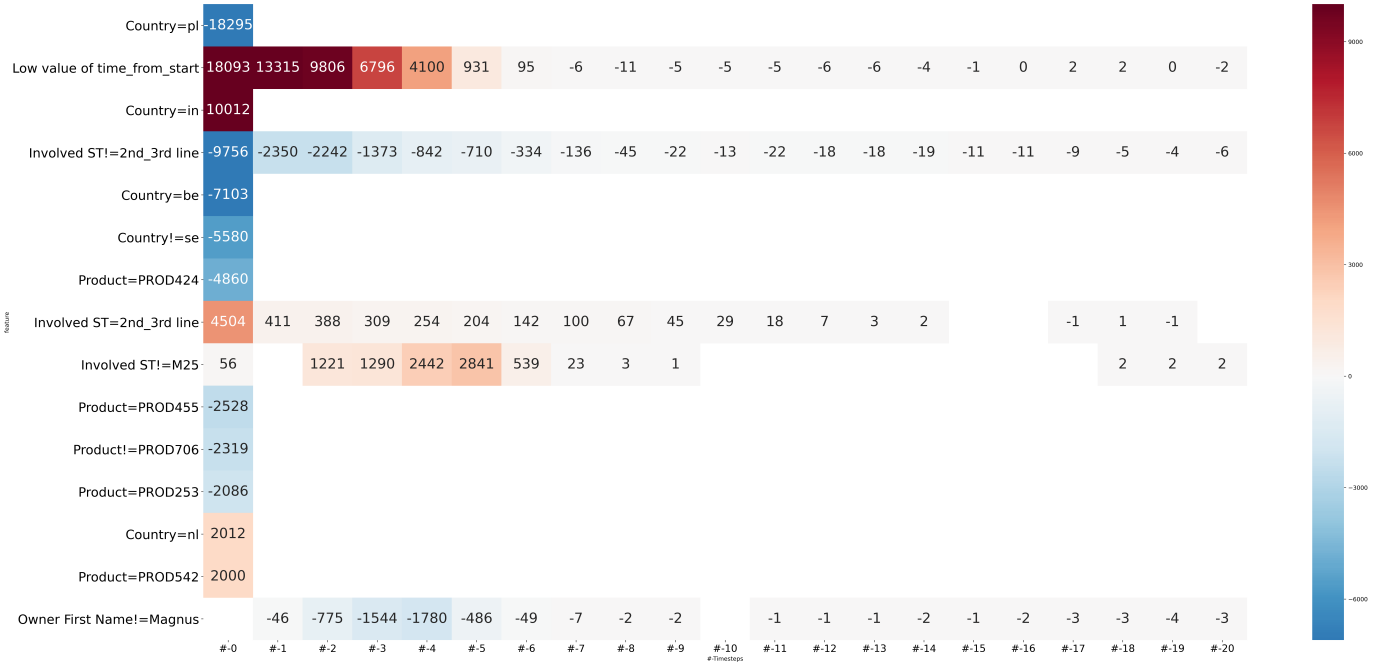

Fig. 9: Offline explanations for BPI13 push to front line prediction

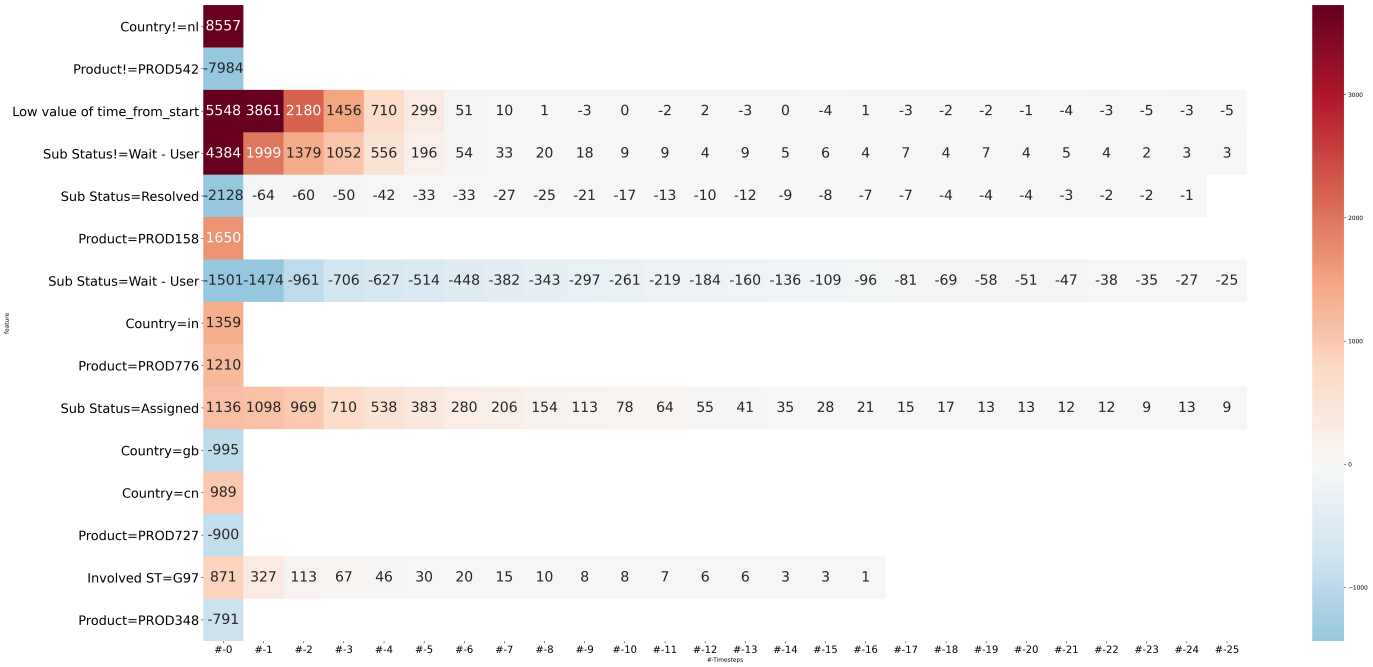

Fig. 10: Offline explanations for BPI13 wait_user prediction

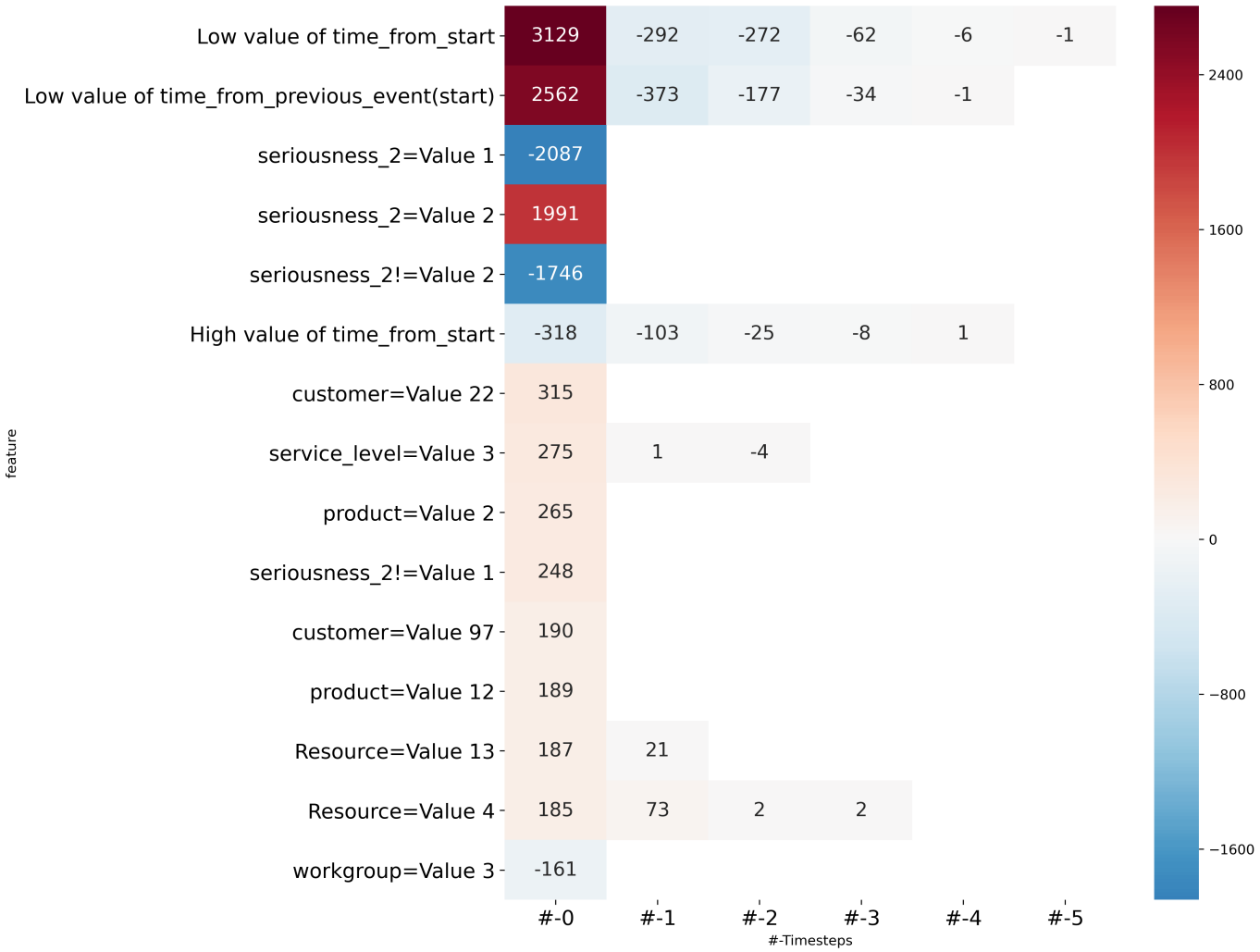

Fig. 11: Offline explanations for HelpDesk17 remaining time prediction