Special Topic in IE - I

Heuristic Graph Prediction Model

Department of Industrial Engineering

Kwon in Yoon

Contents

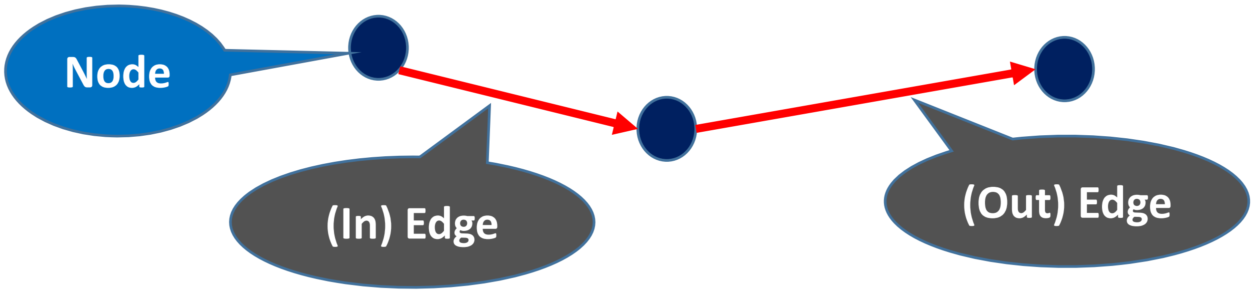
1. Main Idea
2. Prediction Model
3. Experiment
4. Comparison Study
5. Conclusion & Future Work

Heuristic Graph Prediction Model

- The Goal of Heuristic Graph Prediction Model
  - Predict the next activity
- Methodology
  - Graph-based Heuristic Prediction
    - Part 1) Generating the activity graph
    - Part 2) Tracing the activity graph for Prediction

Part 1) Generating the Event Graph

- Generate the Trained Graph
  - Generate graphs of each event through the train data.
  - Generate branches of an event's graph.
    - Trace the entire events in each case.
    - Then, make the list of event flow which is starting the target event of a graph.
  - Take the frequency between events as the label of edge.
  - Convert the frequency into the probability of out edges from a node.



Part 1) Generating the Event Graph

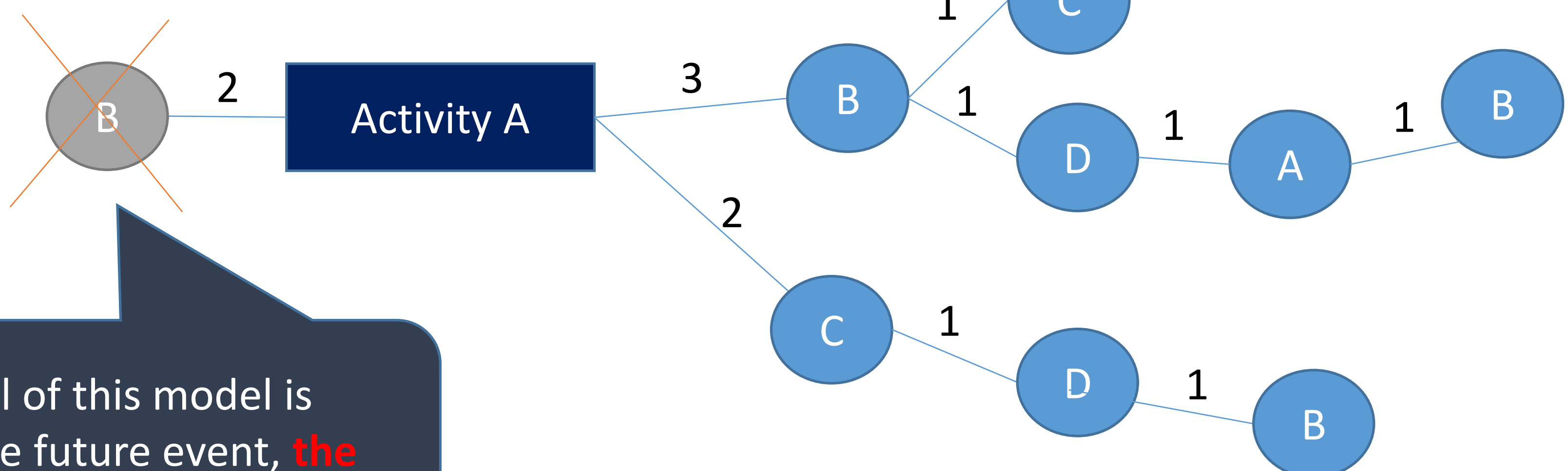
- Example
  - If Creating the graph of an event A
  - The candidate set of nodes
    - Case 1 : [A, B, C]
    - Case 2 : [A, B, D, A, B], [A, B]
    - Case 3 : [A, C, D, B]
    - Case 4 : [A, C]

Case	Event	Case	Event
Case 1	**A**	Case 3	B
	B		**A**
	C		C
Case 2	**A**		D
	B		B
	D	Case 4	B
	A		**A**
	B		C

Part 1) Generating the Event Graph

- Example
- The candidate set of nodes
  - Case 1 : [A, B, C]
  - Case 2 : [A, B, D, A, B], [A, B]
  - Case 3 : [A, C, D, B]
  - Case 4 : [A, C]

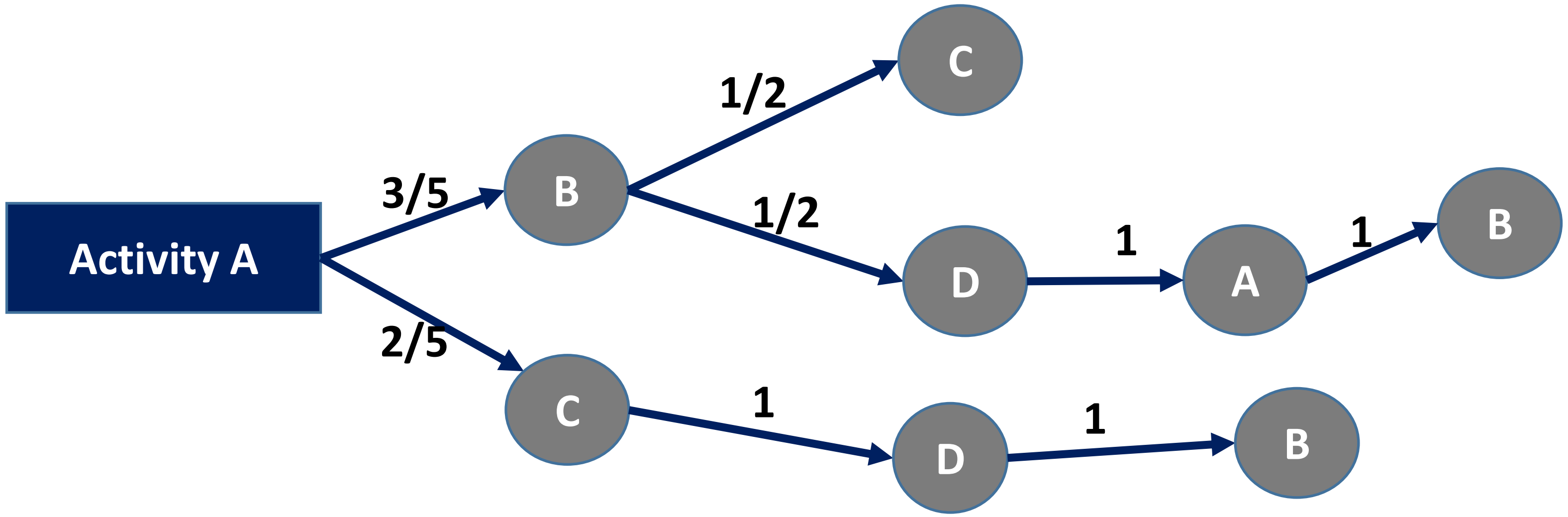
Frequency of the edges
in the candidate sets



As the goal of this model is predicting the future event, **the previous history is no meaningful.**

Part 1) Generating the Event Graph

- The Created Graph
  - Convert the frequency into the probability of out edges from a node.



Part 2) Tracing the Event Graph for Prediction

- Graph Prediction Model

	Act 1	Act2	Act3	The next event
Case	A	B	D	??

1. **Tracing the graph** of the previous events from the latest activity as much as the parameter [Trace_Qty].
2. Take all predicted events **with the probability**.
3. Calculate **the respective score** of the predicted events one by one.
4. Sum the calculated score by each activity.
5. Return the result with the highest score.

Part 2) Tracing the Event Graph for Prediction

- Parameter
  - Trace_Qty (Default : 3)
    - It's for defining how many event graphs would be traced, which is from the last event in each case.
  - Basic Score (Default : 100)
  - Score_Type (Default : 'greedy')
    - ['greedy', 'equally']
  - Score_Rate (Default : 10)
    - It's for defining the score rate to give the different score depending on the distance level.

Part 2) Tracing the Event Graph for Prediction

- Step 1
  - **Tracing the graph** of the previous activities from the latest activity as much as the parameter [Trace_Qty] .

	Event 1	Event 2	Event 3	The next event
Case	**A**	**B**	**D**	**??**

If the parameter [the number of the previous graph to be traced] is defined as 3,

1. Trace the graph A event with B, D event as the next event in an order.
2. Trace the graph B event with D event as the next event.
3. Trace the graph D event.

Part 2) Tracing the Event Graph for Prediction

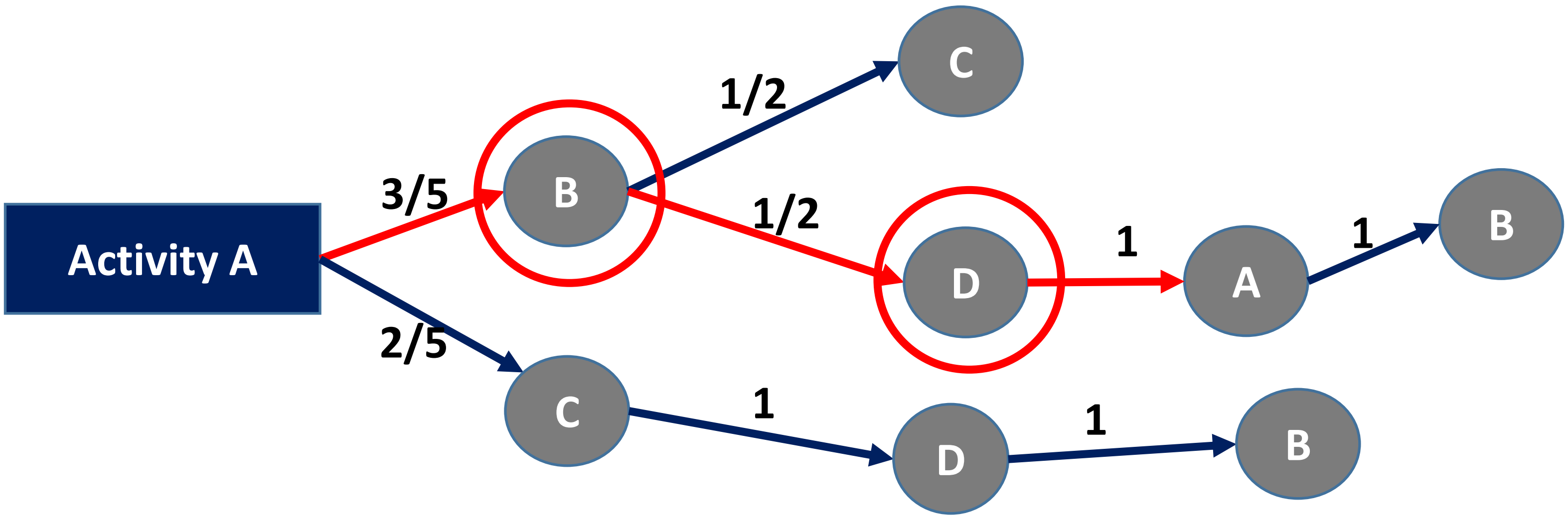
- Step 2
  - Take all predicted activities **with the probability**.
  - Refer to the next page about how to calculate the probability through the graph

	Act 1	Act2	Act3
Case	**A**	**B**	**D**

From	The predicted result	Probability
A graph	C	0.3
B graph	C	0.7
D graph	A	0.2
D graph	B	0.8

Part 2) Tracing the Event Graph for Prediction

- Step 2 (Example)
  - What if tracing the graph A with B, D event as the next event in an order



– $P(B|A,B,D) = P(B|A) * P(D|B) * P(A|D) = \frac{3}{5} * \frac{1}{2} * \frac{1}{1} = \frac{6}{10}$

Part 2) Tracing the Event Graph for Prediction

- Step 3
  - Calculate **the respective score** of the predicted activities one by one.
  - Use the defined parameters (Basic Score = 100, Score_Rate = 10)

From	R	Prob	greedy	equally
A graph (-3)	C	0.3	100*0.3*(10)^2	100*0.3*(10)*3
B graph (-2)	C	0.7	100*0.7*(10)^1	100*0.7*(10)*2
D graph (-1)	A	0.2	100*0.2*(10)^0	100*0.2*(10)*1
D graph (-1)	B	0.8	100*0.8*(10)^0	100*0.8*(10)*1

Part 2) Tracing the Event Graph for Prediction

- Step 3) Score Calculation
  - Option : ‘greedy’, ‘equally_firsthigh’,’ equally_firstlow’
  - ‘greedy’ : Giving higher difference depending on the distance level
    - $Score = Basic\ Score\ * Probability\ * Score\_Rate^{(the\ step\ from\ the\ target\ -1)}$
  - ‘equally_firsthigh’
    - Giving the equal score rate depending on the distance level
    - But, give the higher score for the predicted result from the far event graph.
    - $Score = Basic\ Score\ * Probability * Score\_Rate * (\ Steps\ from\ the\ Target)$

Part 2) Tracing the Event Graph for Prediction

- Step 4, 5
  - Sum the calculated score by each event.
  - Return the result with the highest score.

From	R	Prob	greedy	Equally(10)	Equally(-10)
A graph (-3)	C	0.3	3000	900	-900
B graph (-2)	C	0.7	700	1400	-1400
D graph (-1)	A	0.2	2	200	-200
D graph (-1)	B	0.8	8	800	-800

The predicted activity	greedy	Equally(10)	Equally(-10)
A	20	200	**-200**
B	80	800	-800
C	**3700**	**2300**	-2300

Experiment

- Dataset

  1) Helpdesk1

  – This log records events from a ticketing management system of the help desk of an Italian software company. The log has nine event types (i.e., distinct activities), 3,804 process cases, and 13,710 events.

  2) Hospital Billing

  – This event log records events of the billing of medical services from the financial modules of a regional hospital's ERP system. The event log includes 49,951 events for 10,000 cases, with 16 different event types (distinct activities).

Experiment

- Dataset
  - Train Data / Test Data

Train Data for Graph Generation	Test Data	Test Label
Original Data	Original Data[:-1]	Original Data[-1:]
Prefix 4	Prefix 3	Prefix 4[-1:]
Prefix 5	Prefix 4	Prefix 5[-1:]
Prefix 6	Prefix 5	Prefix 6[-1:]
Prefix 8	Prefix 7	Prefix 8[-1:]
Prefix 11	Prefix 10	Prefix 11[-1:]

Experiment

- Dataset (Test 1)
  - Train Data / Test Data (Prediction by the trained graph from its own data)

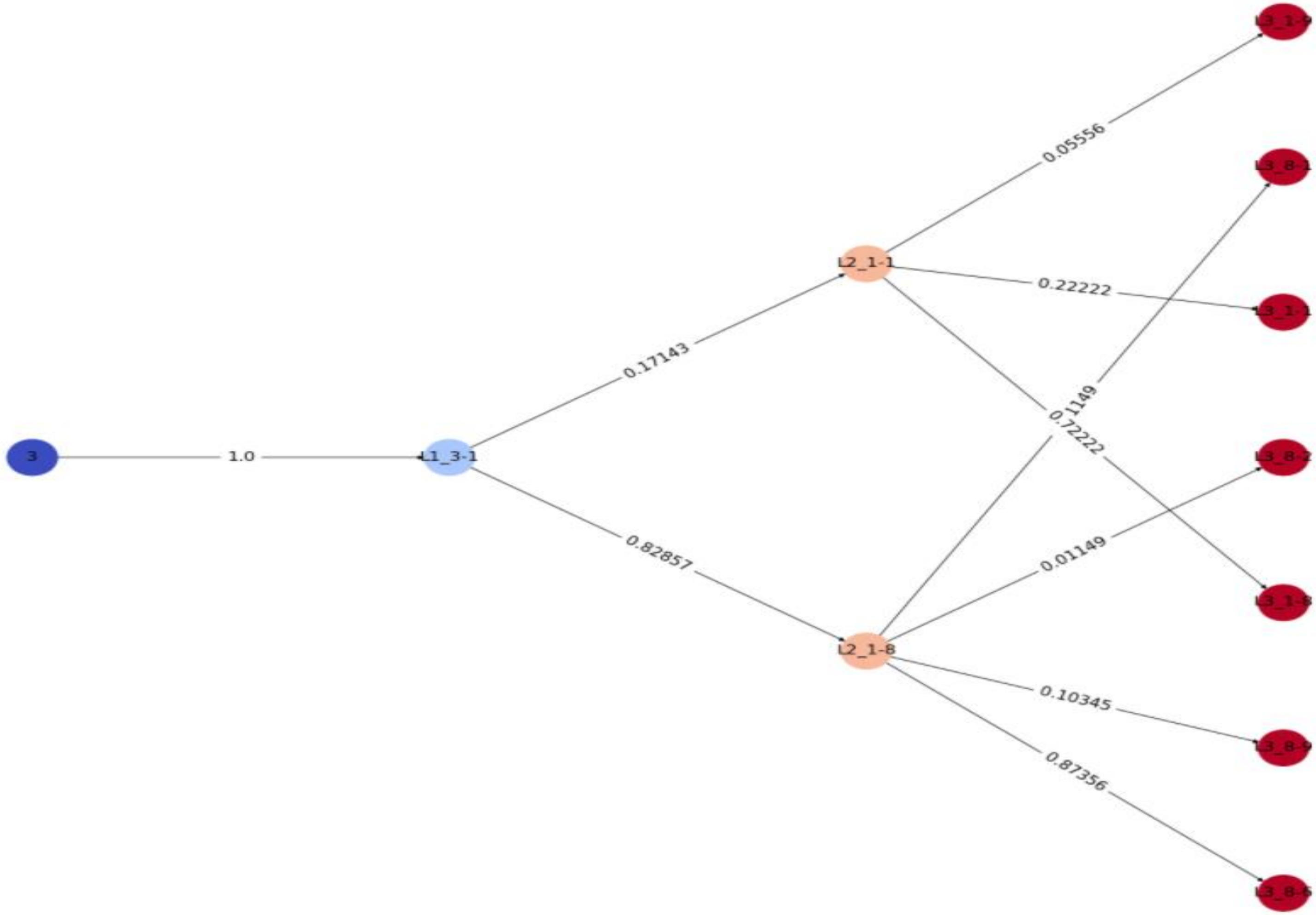
Train Data for Graph Generation	Test Data	Test Label
Original Data	Original Data[:-1]	Original Data[-1:]
Prefix 4	Prefix 3	Prefix 4[-1:]
Prefix 5	Prefix 4	Prefix 5[-1:]
Prefix 6	Prefix 5	Prefix 6[-1:]
Prefix 8	Prefix 7	Prefix 8[-1:]
Prefix 11	Prefix 10	Prefix 11[-1:]

Experiment

- Dataset (Test 2)
  - Train Data / Test Data (Prediction by the trained graph from the original data)

Train Data for Graph Generation	Test Data	Test Label
Original Data	Original Data[:-1]	Original Data[-1:]
	Prefix 3	Prefix 4[-1:]
	Prefix 4	Prefix 5[-1:]
	Prefix 5	Prefix 6[-1:]
	Prefix 7	Prefix 8[-1:]
	Prefix 10	Prefix 11[-1:]

- The Generated Graph



Experiment

- Experiment Result (**helpdesk**)

Score_Type				greedy					equally				
Score_Rate				20	10	1	-10	-20	20	10	1	-10	-20
	Case len	Graph	Trace Qty										
helpdesk_3	2477	Respective	3	37.83	37.83	37.83	2.62	2.62	37.83	37.83	37.83	2.62	2.62
			7	37.83	37.83	37.83	2.62	2.62	37.83	37.83	37.83	2.62	2.62
		Original	3	65.81	65.81	65.81	2.66	2.66	65.81	65.81	65.81	2.66	2.66
			7	65.81	65.81	65.81	2.66	2.66	65.81	65.81	65.81	2.66	2.66
helpdesk_4	1117	Respective	3	36.79	36.79	36.79	6.54	6.54	36.79	36.79	36.79	6.54	6.54
			7	36.79	36.79	36.79	6.54	6.54	36.79	36.79	36.79	6.54	6.54
		Original	3	73.14	73.14	73.14	1.43	1.43	73.14	73.14	73.14	1.43	1.43
			7	73.14	73.14	73.14	1.43	1.43	73.14	73.14	73.14	1.43	1.43
helpdesk_5	490	Respective	3	41.02	41.02	41.02	11.43	11.43	41.02	41.02	41.02	11.43	11.43
			7	41.02	41.02	41.02	11.43	11.43	41.02	41.02	41.02	11.43	11.43
		Original	3	72.24	72.24	72.24	0.82	0.82	72.24	72.24	72.24	0.82	0.82
			7	72.24	72.24	72.24	0.82	0.82	72.24	72.24	72.24	0.82	0.82
helpdesk_7	101	Respective	3	47.52	47.52	47.52	8.91	8.91	47.52	47.52	47.52	8.91	8.91
			7	47.52	47.52	47.52	8.91	8.91	47.52	47.52	47.52	8.91	8.91
		Original	3	10.89	10.89	10.89	4.95	4.95	10.89	10.89	10.89	4.95	4.95
			7	10.89	10.89	10.89	4.95	4.95	10.89	10.89	10.89	4.95	4.95
helpdesk_10	15	Respective	3	53.33	53.33	53.33	20	20	53.33	53.33	53.33	20	20
			7	53.33	53.33	53.33	20	20	53.33	53.33	53.33	20	20
		Original	3	0	0	0	0	0	0	0	0	0	0
			7	0	0	0	0	0	0	0	0	0	0
helpdesk	3803	Original	3	86.17	86.17	86.17	0	0	86.17	86.17	86.17	0	0
			7	86.17	86.17	86.17	0	0	86.17	86.17	86.17	0	0

Experiment

- Experiment Result (**helpdesk**)

Score_Type				greedy					equally				
Score_Rate				20	10	1	-10	-20	20	10	1	-10	-20
	Case len	Graph	Trace Qty										
hospitalbilling_3	24337	Respective	3	63.55	63.55	63.55	0.63	0.63	63.55	63.55	63.55	0.63	0.63
			7	63.55	63.55	63.55	0.63	0.63	63.55	63.55	63.55	0.63	0.63
		Original	3	65.18	65.18	65.18	0.06	0.06	65.18	65.18	65.18	0.06	0.06
			7	65.18	65.18	65.18	0.06	0.06	65.18	65.18	65.18	0.06	0.06
hospitalbilling_4	17066	Respective	3	74.8	74.8	74.8	0.13	0.13	74.8	74.8	74.8	0.13	0.13
			7	74.8	74.8	74.8	0.13	0.13	74.8	74.8	74.8	0.13	0.13
		Original	3	77.12	77.12	77.12	0.06	0.06	77.12	77.12	77.12	0.06	0.06
			7	77.12	77.12	77.12	0.06	0.06	77.12	77.12	77.12	0.06	0.06
hospitalbilling_5	9877	Respective	3	63.15	63.15	63.15	0.19	0.19	63.15	63.15	63.15	0.19	0.19
			7	63.15	63.15	63.15	0.19	0.19	63.15	63.15	63.15	0.19	0.19
		Original	3	67.54	67.54	67.54	0.09	0.09	67.54	67.54	67.54	0.09	0.09
			7	67.54	67.54	67.54	0.09	0.09	67.54	67.54	67.54	0.09	0.09
hospitalbilling_7	3696	Respective	3	33.06	33.06	33.06	0.46	0.46	33.06	33.06	33.06	0.46	0.46
			7	33.06	33.06	33.06	0.46	0.46	33.06	33.06	33.06	0.46	0.46
		Original	3	43.91	43.91	43.91	0.16	0.16	43.91	43.91	43.91	0.16	0.16
			7	43.91	43.91	43.91	0.16	0.16	43.91	43.91	43.91	0.16	0.16
hospitalbilling_10	1631	Respective	3	14.84	14.84	14.84	1.47	1.47	14.84	14.84	14.84	1.47	1.47
			7	14.84	14.84	14.84	1.47	1.47	14.84	14.84	14.84	1.47	1.47
		Original	3	1.9	1.9	1.9	2.33	2.33	1.9	1.9	1.9	2.33	2.33
			7	1.9	1.9	1.9	2.33	2.33	1.9	1.9	1.9	2.33	2.33
hospitalbilling	8163	Original	3	84.23	84.23	84.23	0.04	0.04	84.23	84.23	84.23	0.04	0.04
			7	84.23	84.23	84.23	0.04	0.04	84.23	84.23	84.23	0.04	0.04

- Experiment Result
  - Overall Analysis
    - The greater number of previous events it has, the more accuracy rate the result tends to have.
    - As you can see in the result in the previous page, the result of original data predicted by its own data has the best result.
  - Score Type and Trace Qty
    - There isn't any difference between the results when I set the different score type and trace qty. In the same setting, 2 factors doesn't influence the prediction result considering the accuracy rate.
  - Score rate
    - In case of setting the score rate as positive value, the result (in Red box) was much better than the opposite cases (in Blue box). It conclude that the predicted result from the graph of event which is happened far step ahead from the target is more reasonable.

- Experiment Result (Useful Data only)
  - Remove the same result
  - Take the representative accuracy rate only

	Case len	Graph	Trace Qty	
helpdesk_3	2477	Respective	3	37.83
		Original	3	65.81
helpdesk_4	1117	Respective	3	36.79
		Original	3	73.14
helpdesk_5	490	Respective	3	41.02
		Original	3	72.24
helpdesk_7	101	Respective	3	47.52
		Original	3	10.89
helpdesk_10	15	Respective	3	53.33
		Original	3	0
helpdesk	3803	Original	3	86.17
hospitalbilling_3	24337	Respective	3	63.55
		Original	3	65.18
hospitalbilling_4	17066	Respective	3	74.8
		Original	3	77.12
hospitalbilling_5	9877	Respective	3	63.15
		Original	3	67.54
hospitalbilling_7	3696	Respective	3	33.06
		Original	3	43.91
hospitalbilling_10	1631	Respective	3	14.84
		Original	3	1.9
hospitalbilling	8163	Original	3	84.23

- Experiment Result (helpdesk)

	Case len	Graph	Trace Qty	
helpdesk_3	2477	Respective	3	37.83
		Original	3	65.81
helpdesk_4	1117	Respective	3	36.79
		Original	3	73.14
helpdesk_5	490	Respective	3	41.02
		Original	3	72.24
helpdesk_7	101	Respective	3	47.52
		Original	3	10.89
helpdesk_10	15	Respective	3	53.33
		Original	3	0
helpdesk	3803	Original	3	86.17

Experiment

- Experiment Result (**helpdesk**)
  - Target Graph
    - When the prediction is based on the produced graph by the original data, the result tends to be more accurate.
    - If I pick one interesting point in this perspective, it is the result on helpdesk_7, helpdesk_10. When it predicts the next activity by using the graph trained by its own data, the result was better. In case of helpdesk_10, the accuracy rate of result by the original graph is all 0.
      - I think the reason of it is the lack of test data. The original data has diverse cases, and it created more edges and probability which can cover the overall cases. On the other hands, the test data is too small to judge the accuracy of the model because we can’t guarantee that the test data is normal case.

- Experiment Result (hospitalbilling)

	Case len	Graph	Trace Qty	
hospitalbilling_3	24337	Respective	3	63.55
		Original	3	65.18
hospitalbilling_4	17066	Respective	3	74.8
		Original	3	77.12
hospitalbilling_5	9877	Respective	3	63.15
		Original	3	67.54
hospitalbilling_7	3696	Respective	3	33.06
		Original	3	43.91
hospitalbilling_10	1631	Respective	3	14.84
		Original	3	1.9
hospitalbilling	8163	Original	3	84.23

- Experiment Result (hospitalbilling)
  - When the prediction is based on the produced graph by the original data, the result tends to be more accurate like the result of helpdesk.
  - The difference between helpdesk and hospitalbilling is the comparison result of window size 7.
    - As I guessed, the number of cases in the test data has an effect on the comparison result between by its own graph and the original graph.
      - As the number of cases in hospitalbilling_7 has over 3000 cases, the result of hospitalbilling_7 has the higher accuracy of prediction by the original graph similar to the comparison result of hospitalbilling_3, hospitalbilling_4.

Experiment

- Comparison Study
  - Refer to the result of average accuracy in the paper.
    - Tama, B. A., Comuzzi, M., & Ko, J. (2020). **An Empirical Investigation of Different Classifiers, Encoding, and Ensemble Schemes for Next Event Prediction Using Business Process Event Logs.** *ACM Transactions on Intelligent Systems and Technology (TIST)*, *11*(6), 1-34.
    - Table 11. Results of average accuracy for each classifier and dataset as an individual classifier
  - HP : Heuristic Graph Prediction Model
  - Others : Single Classifier
    - Decision Tree (DT) / Random Tree (RT) / Naive Bayes (NB)
    - Decision Stump (DS) /Support Vector Machine (SVM) / Conjunctive Rule (CR)

- Comparison Study

	Case len	Graph	Trace Qty	HP	DT	C-DT	RT	DS	NB	SVM	k-NN	JRip	OneR	CR	BN	DTNB
helpdesk_3	2477	Respective	3	37.83	66.42	64.92	58.41	62.94	63.22	19.21	58.41	64.03	65.04	62.94	63.59	65.84
		Original	3	65.81												
helpdesk_4	1117	Respective	3	36.79	73.99	73.91	65.82	72.09	70.03	27.84	65.88	73.27	74.12	72.09	69.98	73.27
		Original	3	73.14												
helpdesk_5	490	Respective	3	41.02	74.86	75.34	69.02	72.87	70.64	26.98	68.59	72.87	75.70	72.87	70.23	73.05
		Original	3	72.24												
helpdesk_7	101	Respective	3	47.52	77.89	72.92	70.87	68.55	71.17	35.24	68.55	74.71	74.10	69.73	72.34	75.85
		Original	3	10.89												
helpdesk	3803	Original	3	86.17												

- The result of the original data has the best accuracy rate among all result.
- In case of window 3,4,5 by the original graph, those have the result comparable to the result of superior classifiers.
- In case of window 7, the result was the worst among all classifier results.

Experiment

- Comparison Study

	Case len	Graph	Trace Qty	HP	DT	C-DT	RT	DS	NB	SVM	k-NN	JRip	OneR	CR	BN	DTNB
hospitalbilling_3	24337	Respective	3	63.55	92.63	92.58	88.10	61.43	90.98	59.07	88.10	92.61	92.27	61.43	90.93	92.51
		Original	3	65.18												
hospitalbilling_4	17066	Respective	3	74.8	91.23	91.20	86.36	73.42	87.97	47.80	86.34	90.96	90.56	73.42	88.07	91.01
		Original	3	77.12												
hospitalbilling_5	9877	Respective	3	63.15	88.31	88.33	83.05	58.26	83.84	30.60	82.97	87.91	87.45	58.26	83.95	87.76
		Original	3	67.54												
hospitalbilling_7	3696	Respective	3	33.06	85.41	85.05	81.03	48.94	77.63	30.17	80.71	84.68	84.27	48.94	78.19	84.03
		Original	3	43.91												
hospitalbilling	8163	Original	3	84.23												

  - HP got the bed accuracy comparing to the previous result.
    - In case of prefix 7, the accuracy was around 40 %.
    - HP has the similar tendency of the accuracy rates to DS, CR.

Conclusion & Future Work

- Conclusion
  - HP model itself has the similar result regardless of data.
  - The longer edges the trained graph has, the better the result is.
  - In case of dataset ‘hospitalbilling’,
    - The result of other classifiers have the significantly good result. On the other hands, HP, DS, SVM, CR have the relatively bed result.
    - Especially, HP, DS, CR have the similar result depending on the prefix size.

Conclusion & Future Work

- Future Work
  - Apply GNN (Graph Neural Network) and GATs (graph attention networks)
    - By using the other static information of events or other attributes in cases, the graphs can be applied to other networks to enhance the performance.
    - Through that, it's possible to extract the more importance information of events to predict the next event.
  - Study Decision Stump (DS) / Conjunctive Rule (CR) which have the similar result
    - Analyze what kind of factor made the similar result

Thank you.

ULSAN NATIONAL INSTITUTE OF SCIENCE AND TECHNOLOGY