Chapter 4

Path Weight Aggregation Feature for Link Prediction in Dynamic Networks (PWAF)

The goal of this chapter is to propose Path Weight Aggregation Feature (PWAF), which is a new feature based on ranking multi edge occurrences across the entire network. Different topological aspects of the networks (Local, Global, and Quasi-local) as well as Clustering Coefficient based features are taken into consideration for feature generation, in addition to the suggested Path Weight-Based Aggregation Feature (PWAF). One of the features used for better prediction is the Level-2 node clustering coefficient (CCLP2). This chapter [1] studies a new path weight based feature of social networks.

[1]Published in Computer communication PWAF : Path Weight Aggregation Feature for Link Prediction in Dynamic Networks

4.1 Introduction

One of the most basic challenges in complex network analysis is the link prediction problem. Link prediction is useful in a wide variety of disciplines. In the Internet and web science domains, these topics include automatic hyperlink construction [27], website hyper-link prediction [28], and recommendation of friend system on online social networks like Facebook and Instagram [136]. Combinations of different social networks have been also explored as link prediction in multiplex networks [137, 138]. Protein-protein interactions (PPI) and bio-informatics have also assisted from link prediction [31]. Link prediction can be used in the world of security to uncover hidden links between terrorists and their organization. For various sorts of graphs, researchers have used a variety of link prediction techniques. There are several sorts of these strategies, including similarity-based, probabilistic, and learning-based models [1, 139–144] etc. We deal with similarity-based indices in this paper, which cover Local, Global, and Quasi-local indices..

The local similarity scores are obtained using information extracted from the local neighborhood of nodes. Examples of such indices are Common Neighbors (CN), Adamic/Adar Index (AA), Jaccard Coefficient (JC), Preferential Attachment (PA).

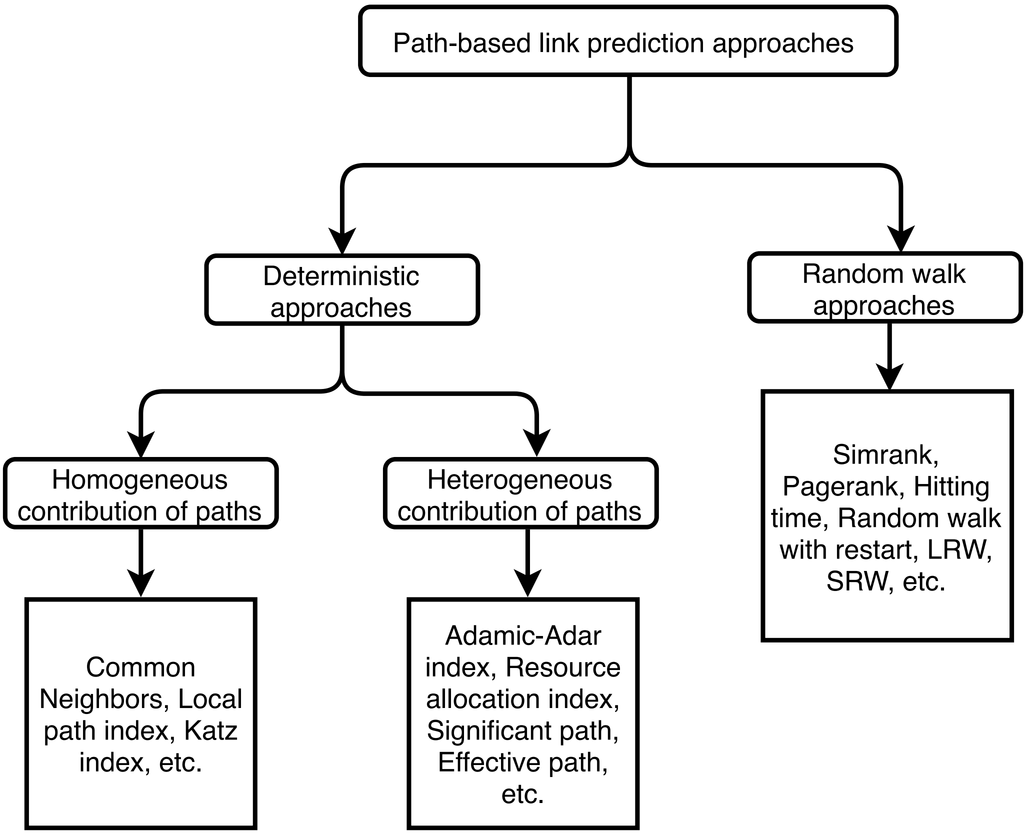
The whole topological information of a network is used to create global similarity-based indices. These methods are more computationally intensive than local similarity-based methods, but they provide a more comprehensive view of the graph structure. Global similarity-based indices includes shortest path (SP), COS+ (COSP), Matrix Forest Index (MFI), and Average Commute Time (ACT).

To strike a compromise between the aspects of local and global similarity-based measures, quasi-local similarity-based techniques have been implemented. Quasi-local similarity-based indices includes local path Index (LP), Path of Length 3 (L3), Local Random Walk (LRW) and Superposed Random Walk (SRW). Quasi-local link prediction algorithms are based on random walks [145, 146].

To make the rich feature-set for better prediction we have also taken different clustering algorithms. The clustering algorithms considered are Clustering Coefficient based Link Prediction (CCLP), Node and Link clustering coefficient (NLC) , CAR-based Common Neighbor Index (CARCN) . We have also used Level-2 node clustering coefficient (CCLP2) for more accurate prediction. These clustering approaches provide more information about the nodes and edges, improving accuracy.

The existing works can be classified into the following taxonomy from the perspective of pathways shown in Figure 4.1.

FIGURE 4.1: Path-based approaches to link prediction [2]



The following are the key motivation for developing this approach.

- According to Ajay et al. [1], link prediction accuracy improves with more local, global, quasi-local, and clustering topology information. We extracted information from the local, global, and quasi-local as well as clustering information from multi hop neighborhoods to achieve this effect in our method.

- In this proposal, we have computed clustering coefficient of Level-1 and Level-2 common neighbor [36] of the seed node pair for more information for better accuracy. In comparison to the Level-1 frequent neighbors and their corresponding clustering coefficients, Level-2 explores more information about networks [147].

- The proposed method employs feature vectors of node pairs generated from all snapshots to incorporate several types of structural information (topological based which includes Local, Global and Quasi-local similarity, clustering coefficient based algorithm, and PWAF-based).

The following are the major contributions of this paper:

- In this paper, we present a new Path Weight Aggregation Feature (PWAF) feature to address link prediction in dynamic networks.

- In our work, we offer a link prediction framework that uses various semantic topological features (Local, Global and Quasi-local similarity), clustering features like CCLP, NLC, CARCN and Level-2 node clustering coefficient with different machine learning models.

- We tested individual traditional link prediction approaches in several machine learning models with our proposed PWAF model after synthesizing these feature sets.

- Also, we have compared the *PWAF* machine learning variations PWAF-NN, PWAF-LR, PWAF-XGB, PWAF-RFC and PWAF-LDA with state-of-the-art link prediction algorithms using the five performance evaluation metric. We have observed a significant increase in performance based on the results of these metric.

We deployed a variety of machine learning methods on real-world dynamic datasets to test our methodology. According to Memon et al. [148]., there are various advantages to using various machine learning algorithms over traditional methods. After reviewing the data, we discovered that our method produces better results. On seven different open dynamic datasets with five performance assessment parameters, we compared our performance against five state-of-the-art methods [131–133, 147, 149]. Experiments show that our proposed strategy improves performance significantly.

4.2 Proposed work

The majority of recent research uses network topology to extract feature sets. These characteristics are generic and domain-independent, and they can be used in any network [131, 150, 151]. Other research focuses on identifying node and edge information that are critical for improving link prediction performance. Typical, neighbourhood, and path-based features are examples of such features [152, 153]. Some related literature also shows that the clustering coefficient is related to the link prediction problem [154]. The link prediction problem is described as a binary classification problem. The class label is indicated by the presence or absence of connections. If a link exists between two nodes, the label is set to 1; otherwise, the label is set to 0. Link prediction model using machine learning techniques in dynamic networks is shown in Fig. 4.4

4.2.1 Path Weight Aggregation Feature $(PWAF)$ for link prediction

It is based on weight of the paths between two nodes $i\&j$. Since we are creating the graphs from snapshots, we define the weight of each occurring edge in a snapshot as the number of times that edge is encountered in a given snapshot. This gives us the relative significance of edges compared to each other. The PWAF feature is mathematically computed by the following equations (here $SOP(i, j)$ is the sum of common neighbor

FIGURE 4.2: Graph of F($\frac{W}{\sqrt{N}}$) Vs $\frac{W}{\sqrt{N}}$ graph

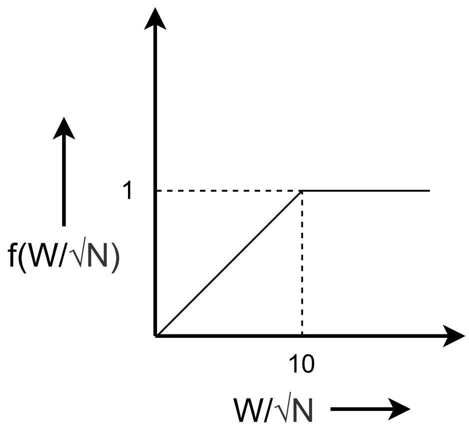
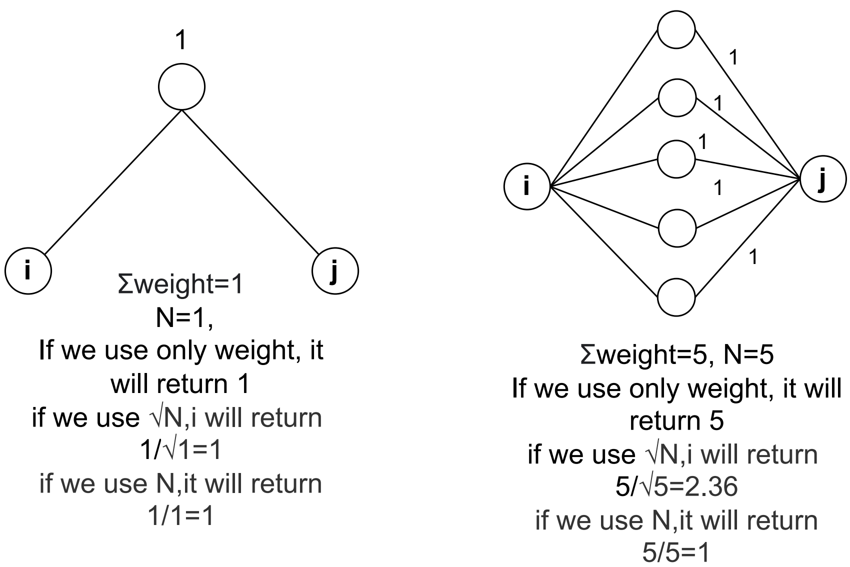


FIGURE 4.3: Analysis for selecting $\sqrt{N}$



paths between nodes $i\&j$ and we are calculating PWAF feature value for one specific snapshot):

$$SOP(i,j) = \sum_{t=1}^{N} wtgf(i,t) + wtgf(t,j) \tag{4.1}$$

$$PWAF(i,j) = \begin{cases} 1 & x_{ij} < 2 \\ \frac{F(SOP(i,j))}{2\times\sqrt{N}} & x_{ij} = 2 \\ 0 & x_{ij} > 2 \end{cases} \tag{4.2}$$

$$F(x) = \begin{cases} \frac{x}{10} & x \leq 10 \\ 1 & x \geq 10 \end{cases} \tag{4.3}$$

where x_{ij} is the distance between nodes and N is common neighbor nodes between $i\&j$. The detailed process for the calculation of weight is given in the proposed algorithm Algorithm 2.

4.2.2 Analysis for selecting $\sqrt{N}$

.

The Fig. 4.2 shows the graph of F($\frac{W}{\sqrt{N}}$) Vs $\frac{W}{\sqrt{N}}$.

Here, when x is less than equal to 10, it is a straight line with slope $\frac{1}{10}$. when x is greater than 10 then the value of the function is constant that is 1, which is parallel to x-axis. In this case $x = \frac{w}{\sqrt{N}}$ which is calculated using the summation of weight and number of links [155]. Here, w is equal to $SOP(i,j)$ (Equation 4.1). The example is shown in Fig. 4.3.

4.2.3 Proposed PWAF Feature Generation Algorithm

The detailed algorithm of Path Weight-Based Aggregation Feature (PWAF) is shown in Algorithm 2. In this algorithm, the line number 1-2 is the initialization phase. We have initializes *num_indirect* and *wt_indirect* with value 0. We have taken variables *wty* and *numy* to store weight and number of links respectively. Lets consider an edge between nodes $x\&y$ ($e[0]\&e[1]$ in algorithm) for prediction. The for loop in line number 6 is

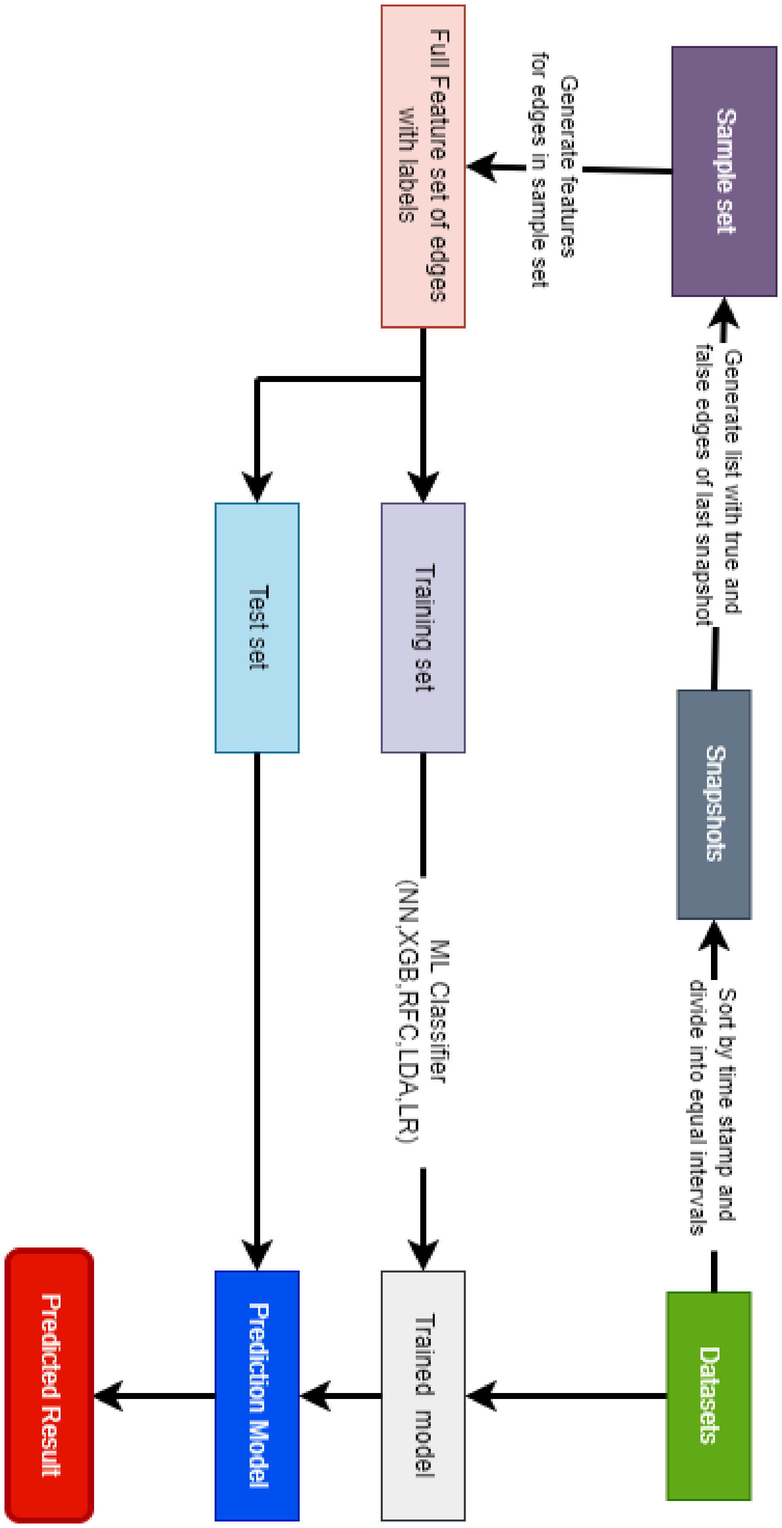


FIGURE 4.4: Representations of link prediction model using machine learning techniques in dynamic networks

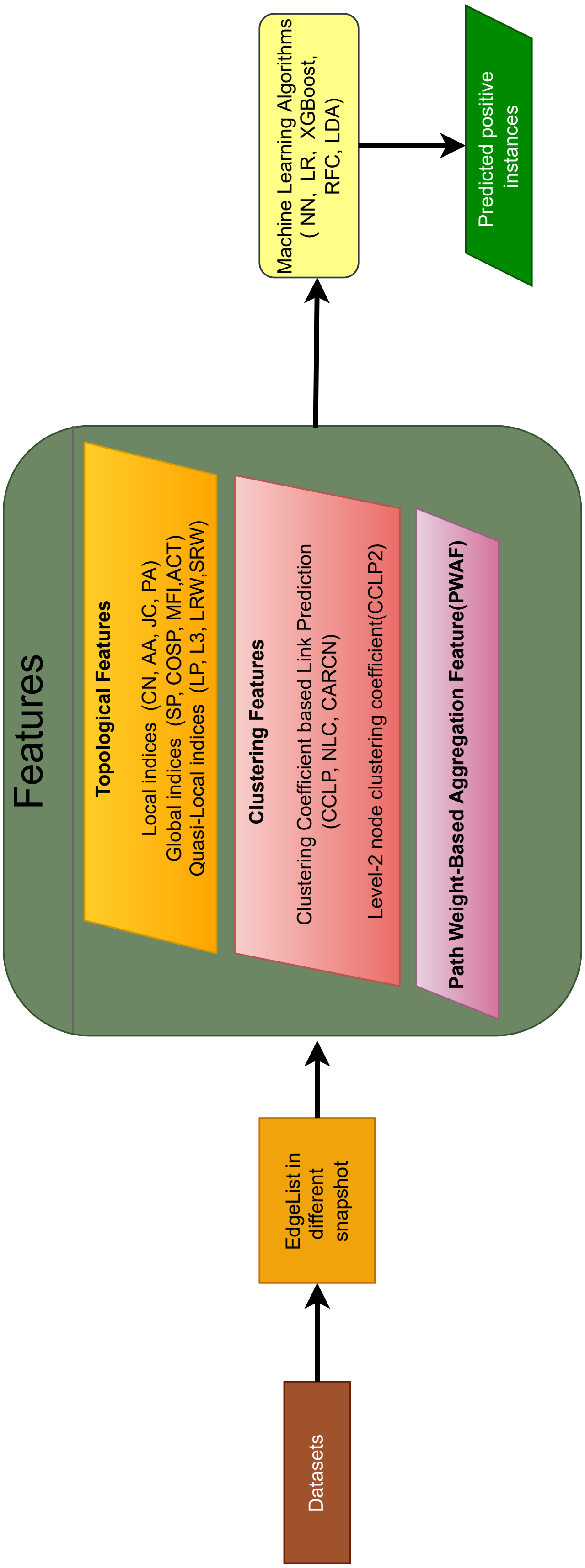


FIGURE 4.5: PWAF model overview

Algorithm 2: Path Weight-Based Aggregation Feature (PWAF) algorithm

Input: $G(V,E)$: Dynamic social graph, v: no. of nodes, e: edge for which the value is to be calculated, $wtgf$: frequency of edge occurence, t= current snapshot

Output: Return feature value

```
1  num_indirect ← 0                                          ▷ Initialization Phase
2  wt_indirect ← 0
3  if [e[0], e[1]] ∈ E then
4      return 1
5  else
6      for each i ∈ n do
7                                  ▷ if i is equal to any node of edge e then Pass
8          if i == e[0] —— i == e[1] then
9              Pass
10                                 ▷ if node i acts as a common neighbor then
11             ▷ add weight of both edge to wt_indirect and increment num_indirect
12         if (wtgf[t][e[0]][i] ≠ 0) & (wtgf[t][i][e[1]] ≠ 0) then
13             num_indirect = num_indirect + 2
14             wt_indirect = wt_indirect + wtgf[t][e[0]][i] + wtgf[t][i][e[1]]
15     if wt_indirect == 0 then
16         return 0
17     else
18         PWAF = wt_indirect / (10 × √num_indirect)                 ▷ Feature value
19         if PWAF > 1 then
20             PWAF = 1          ▷ Normalizing exceptional cases using Eq. 4.3
21         return PWAF
```

for iteration among all nodes. Line 6-12 is used for iteration over all possible common neighbors to update values of variables *num_indirect* and *wt_indirect*. In Line 7-8, if i is equal to any node of edge e then it will it go for next iteration. In line 9-11, it is checked if intermediate node i is common neighbor of $x\&y$ in that particular snapshot. If considered node i is an intermediate node i.e., if edges $[x,i]$ and $[i,y]$ are present, add weight of edges $[x,i]$ and $[i,y]$ to *wt_indirect* and increase *num_indirect* by 2. In lines 13-14, if the value of *wt_indirect* is zero it will return zero. In lines 16-19, we will get the feature value. Here we will normalize for some special cases and then we put these values of *wt_indirect* and *num_indirect* in the formula given in Equation 4.2. The formula is used to calculate the feature value of all the links in a particular snapshot which is further used for link prediction. An important point to note in the calculation procedure for our proposed feature *PWAF* is that the added computation is mostly carried out at the time of snapshot creation where we define the weight of each edge as the number of times the edge has occurred in that particular snapshot. Since the whole edge list of dynamic network would have to be compulsorily read for time based edge sorting and snapshot graph creation, this does not constitute much added complexity. After this weight dictionary has been calculated for each snapshot at the time of snapshot creation, the individual *PWAF* features calculated for every snapshot have a maximum complexity of $\mathcal{O}(V * D)$, for lines 6-12, where we iterate over the whole list of nodes and check for the existence of indirect paths. Here D is the average degree of graph. After the variables *wt_indirect* and *num_indirect* have been computed in these lines, the calculation of final *PWAF* feature value using line 16 is trivial.

The overall process of link prediction using our PWAF model is depicted in Figure 4.5.

4.3 Result Analysis

All of the experiments in this study were carried out using a system with an AMD Ryzen 2700 8-core processor, 32 GB of DDR4 RAM running at 2666 MHz, and a 512 GB NVME SSD hard drive. Python version 3.6 was used to programme. The experimental results obtained from the experiments are examined in this section. To begin, we compare the PWAF result to individual features in different machine learning algorithms such as Neural Network (NN), Logistic Regression (LR), XGBoost(XGB), Random Forest Classifier (RFC), and Linear Discriminant Analysis (LDA) using five performance evaluation metrics: AUPR, F1 score, Average Precision, Balance Accuracy score, and AUC on seven well known temporal datasets. We have also compared the performance of PWAF different machine learning variations to five state-of-the-art algorithms, RA [149], PROXM [133], WEAK [131], Node2Vec [132] and LGQ [147]. On seven

TABLE 4.1: Performance of PWAF model and its machine learning variations- XGBoost (XGB)

	Dataset	CN	AA	JC	PA	SP	COSP	ACT	MFI	LOCALP	L3	SRW	LRW	CCLP	CCLP2	NLC	CARCN	PWAF-XGB
AUPR	mit	0.62625	0.5984	0.58922	0.68087	0.74597	**0.76213**	0.68131	0.71961	0.66398	0.68837	0.48126	0.49856	0.63603	0.67342	0.64951	0.63686	0.75111
	radoslaw-email	0.59358	0.59676	0.6178	0.49785	**0.77395**	0.76579	0.54894	0.75489	0.57371	0.58737	0.32404	0.37342	0.61116	0.59208	0.60544	0.59083	0.76699
	Eu-core	0.69628	0.71029	0.70206	0.45924	**0.87151**	0.8311	0.46734	0.76396	0.68993	0.73754	0.29119	0.34845	0.73703	0.72534	0.74635	0.70035	0.86499
	fb-forum	0.34143	0.35031	0.39454	0.44669	0.82946	0.8289	0.52051	0.79621	0.63642	0.82522	0.30137	0.32461	0.36985	0.40569	0.41904	0.35738	**0.85159**
	CollegeMsg	0.29127	0.21793	0.32036	0.44565	**0.62995**	0.59764	0.45002	0.56692	0.49475	0.58364	0.36212	0.3026	0.27928	0.35197	0.34531	0.27329	0.61114
	mathoverflow	0.63681	0.64311	0.63881	0.64687	0.64303	0.65138	0.61731	0.62227	0.65352	0.66016	0.53304	0.53148	0.63968	0.65317	0.64812	0.6405	**0.67461**
	lkml-reply	0.71726	0.72167	0.70166	0.71247	0.69959	0.71539	0.68318	0.696	0.7179	0.72815	0.62095	0.62054	0.7208	0.72405	0.72452	0.71735	**0.74067**
F1 score	mit	0.56838	0.54324	0.51847	0.63305	0.71601	**0.72991**	0.63545	0.68334	0.61812	0.64368	0.34957	0.37972	0.58302	0.62974	0.59676	0.58582	**0.72016**
	radoslaw-email	0.51765	0.52235	0.55223	0.37559	0.71698	**0.7241**	0.43997	0.70861	0.4818	0.5002	0.0427	0.03924	0.55065	0.4987	0.54186	0.51803	**0.72509**
	Eu-core	0.66253	0.67837	0.66963	0.29696	**.86012**	0.81683	0.32399	0.74311	0.65348	0.71129	0.05187	0.07776	0.70909	0.69195	0.716	0.66643	0.85368
	fb-forum	0.07238	0.09341	0.12852	0.29119	0.80139	0.80503	0.36616	0.76881	0.58424	0.79708	0.09336	0.10468	0.10671	0.15026	0.16183	0.05801	**0.82989**
	CollegeMsg	0.06685	0.03379	0.06042	0.26071	0.41101	0.41936	0.22975	0.39494	0.32778	0.42316	0.1236	0.10808	0.05632	0.11334	0.09712	0.0626	**0.44215**
	mathoverflow	0.52059	0.51092	0.51917	0.53083	0.51652	0.51956	0.47025	0.5138	0.52756	0.54341	0.42392	0.4236	0.50858	0.51576	0.50815	0.45126	**0.55717**
	lkml-reply	0.64299	0.64608	0.66024	0.64011	0.65215	0.64267	0.60047	0.62407	0.64886	0.65716	0.55451	0.55341	0.64254	0.65391	0.65259	0.64111	**0.67096**
AVG PRECISION	mit	0.43189	0.40477	0.3995	0.49417	0.57059	**0.59555**	0.49373	0.53689	0.47167	0.50272	0.30853	0.32001	0.44177	0.48267	0.45797	0.44219	0.57971
	radoslaw-email	0.39181	0.39434	0.41636	0.30492	**0.60507**	0.59845	0.34618	0.58274	0.37034	0.3841	0.20442	0.20853	0.41114	0.38695	0.40366	0.38911	0.59996
	Eu-core	0.50357	0.52171	0.51037	0.2527	**0.76245**	0.69434	0.25909	0.59179	0.49577	0.55817	0.16878	0.16753	0.55674	0.54151	0.5702	0.50917	0.75083
	fb-forum	0.1839	0.18615	0.19726	0.24879	0.69112	0.69173	0.29427	0.64157	0.42956	0.68444	0.17814	0.18448	0.19339	0.20338	0.20971	0.17969	**0.72856**
	CollegeMsg	0.17282	0.17631	0.17794	0.23837	0.35786	0.3465	0.23385	0.32435	0.27249	0.34282	0.19398	0.18005	0.18049	0.1913	0.18798	0.16913	**0.36126**
	mathoverflow	0.40686	0.40622	0.40729	0.41703	0.40872	0.41473	0.37493	0.39524	0.41976	0.43063	0.31553	0.31465	0.40367	0.41439	0.40813	0.37892	**0.44589**
	lkml-reply	0.51497	0.51994	0.50839	0.50985	0.50399	0.51354	0.47076	0.4902	0.51818	0.53019	0.40881	0.40843	0.51764	0.52531	0.52488	0.51451	**0.54658**
BAL ACC SCORE	mit	0.7201	0.70537	0.68384	0.75841	0.82573	**0.82741**	0.76179	0.801	0.75422	0.76724	0.60043	0.61463	0.73124	0.76367	0.73509	0.73366	**0.82597**
	radoslaw-email	0.68733	0.69024	0.70848	0.61422	0.79646	**0.80967**	0.64458	0.79854	0.66633	0.67596	0.50782	0.50764	0.709	0.67438	0.70409	0.68861	**0.8101**
	Eu-core	0.78839	0.79836	0.79309	0.58579	**0.92026**	0.90082	0.59705	0.85635	0.77917	0.82304	0.50112	0.50117	0.81752	0.80114	0.81715	0.78984	**0.92083**
	fb-forum	0.51547	0.52021	0.5306	0.58274	**0.85933**	**0.86826**	0.61475	0.84618	0.73325	0.85692	0.51857	0.52271	0.52392	0.53684	0.54035	0.51244	**0.88229**
	CollegeMsg	0.50374	0.50351	0.51192	0.57215	0.63147	0.63603	0.56265	0.62596	0.59889	0.63791	0.5285	0.52208	0.50991	0.52519	0.5216	0.50302	**0.64581**
	mathoverflow	0.68422	0.67794	0.68313	0.68899	0.68108	0.68188	0.65861	0.68219	0.68614	0.69463	0.64086	0.64083	0.67697	0.67954	0.67586	0.64798	**0.70083**
	lkml-reply	0.75157	0.75265	**0.77709**	0.75085	0.76879	0.75196	0.72849	0.74324	0.75632	0.75963	0.71133	0.71043	0.75001	0.7584	0.75704	0.74997	0.76681
AUC	mit	0.7201	0.70537	0.68384	0.75841	**0.82573**	0.82741	0.76179	0.801	0.75422	0.76724	0.60043	0.61463	0.73124	0.76367	0.73509	0.73366	**0.82597**
	radoslaw-email	0.68733	0.69024	0.70848	0.61422	0.79646	0.80967	0.64458	0.79854	0.66633	0.67596	0.50782	0.50764	0.709	0.67438	0.70409	0.68861	**0.8101**
	Eu-core	0.78839	0.79836	0.79309	0.58579	**0.92026**	0.90082	0.59705	0.85635	0.77917	0.82304	0.50112	0.50117	0.81752	0.80114	0.81715	0.78984	**0.92083**
	fb-forum	0.51547	0.52021	0.5306	0.58274	0.85933	0.86826	0.61475	0.84618	0.73325	0.85692	0.51857	0.52271	0.52392	0.53684	0.54035	0.51244	**0.88229**
	CollegeMsg	0.50374	0.50351	0.51192	0.57215	0.63147	0.63603	0.56265	0.62596	0.59889	0.63791	0.5285	0.52208	0.50991	0.52519	0.5216	0.50302	**0.64581**
	mathoverflow	0.68422	0.67794	0.68313	0.68899	0.68108	0.68188	0.65861	0.68219	0.68614	0.69463	0.64086	0.64083	0.67697	0.67954	0.67586	0.64798	**0.70083**
	lkml-reply	0.75157	0.75265	0.77709	0.75085	**0.76879**	0.75196	0.72849	0.74324	0.75632	0.75963	0.71133	0.71043	0.75001	0.7584	0.75704	0.74997	**0.76681**

TABLE 4.2: Performance of PWAF model and its machine learning variations- Random Forest Classifier (RFC)

	Dataset	CN	AA	JC	PA	SP	COSP	ACT	MFI	LOCALP	L3	SRW	LRW	CCLP	CCLP2	NLC	CARCN	PWAF-RFC
	mit	0.55936	0.5916	0.53592	0.67012	0.70546	0.71285	0.68237	0.71668	0.65341	0.66824	0.47016	0.4788	0.56775	0.65203	0.60391	0.58651	**0.78**
	radoslaw-email	0.52742	0.59078	0.58602	0.5003	0.80918	0.79987	0.54783	0.79148	0.56092	0.58373	0.412	0.40654	0.59529	0.57766	0.60115	0.56907	**0.81991**
	Eu-core	0.65703	0.72692	0.70351	0.39398	0.86129	0.86251	0.43611	0.80594	0.67968	0.76218	0.23019	0.24003	0.73735	0.72628	0.75203	0.68188	**0.90094**
AUPR	fb-forum	0.24188	0.22507	0.25416	0.40396	0.84152	0.81633	0.49603	0.83065	0.65634	0.82422	0.25828	0.24996	0.24954	0.26328	0.26712	0.28978	**0.86375**
	CollegeMsg	0.20956	0.20554	0.21106	0.29562	0.49765	0.46035	0.31815	0.45151	0.36595	0.48568	0.27387	0.25489	0.20686	0.24648	0.2353	0.22715	**0.55706**
	mathoverflow	0.62268	0.61983	0.61349	0.59191	0.59564	0.59717	0.54289	0.55275	0.60027	0.59898	0.50971	0.50855	0.61337	0.61695	0.62308	0.6214	**0.63686**
	lkml-reply	0.72107	0.71577	0.69933	0.69146	0.70169	0.67772	0.65904	0.67497	0.70306	0.69519	0.60386	0.60161	0.71515	0.70405	0.72432	0.71974	**0.7399**
	mit	0.58805	0.61386	0.54797	0.64533	0.7327	0.72207	0.67485	0.71172	0.6245	0.65409	0.51504	0.52652	0.60118	0.6695	0.61931	0.61404	**0.75758**
	radoslaw-email	0.55221	0.56869	0.5684	0.52001	0.67829	0.68737	0.53934	0.67687	0.56696	0.58684	0.46747	0.47035	0.5663	0.57755	0.57157	0.56423	**0.69715**
	Eu-core	0.63776	0.65758	0.64173	0.41488	**0.84103**	0.80931	0.42721	0.74663	0.62985	0.70037	0.31021	0.31616	0.67849	0.65569	0.67114	0.65138	0.82693
F1 score	fb-forum	0.27366	0.25717	0.28058	0.39608	**0.80374**	0.7644	0.46846	0.74497	0.59296	0.71723	0.32805	0.32325	0.28911	0.3	0.29926	0.1951	0.78619
	CollegeMsg	0.20767	0.19409	0.20915	0.3162	0.39318	0.39817	0.33461	0.38206	0.36668	0.41935	0.29325	0.28905	0.18013	0.21383	0.19172	0.11637	**0.46203**
	mathoverflow	0.52992	0.52664	0.52833	0.53741	0.51474	0.52575	0.50237	0.51241	0.54074	0.5385	0.51333	0.51171	0.5159	0.52293	0.51489	0.46328	**0.55742**
	lkml-reply	0.6461	**0.65174**	**0.65348**	0.63485	**0.65215**	0.63164	0.61331	0.61257	0.64239	0.64146	0.60384	0.60362	**0.65169**	0.64967	**0.65657**	0.64063	**0.65899**
	mit	0.56331	0.59347	0.53874	0.67025	0.68212	0.71229	0.68314	0.71999	0.65399	0.6704	0.4736	0.48186	0.57102	0.65673	0.60746	0.59008	**0.78113**
	radoslaw-email	0.52657	0.58914	0.5841	0.4989	0.77613	0.79572	0.54554	0.78688	0.55894	0.58217	0.41098	0.40585	0.59298	0.57584	0.59832	0.56771	**0.81697**
	Eu-core	0.65644	0.72429	0.70089	0.39275	0.87395	0.85825	0.43438	0.80312	0.67659	0.75879	0.23163	0.24089	0.7344	0.72336	0.74865	0.67695	**0.89868**
AVG PRECISION	fb-forum	0.2328	0.21536	0.24276	0.40347	0.83448	0.80091	0.49453	0.82533	0.65253	0.81675	0.26029	0.25238	0.22984	0.23971	0.23834	0.20973	**0.85941**
	CollegeMsg	0.19621	0.19471	0.19169	0.2909	0.47723	0.41648	0.31251	0.4328	0.35212	0.46043	0.2704	0.25298	0.19291	0.21123	0.20072	0.18132	**0.53456**
	mathoverflow	0.48461	0.47182	0.46779	0.5163	0.52288	0.47411	0.46548	0.49382	0.52605	0.50197	0.44357	0.44227	0.46053	0.45885	0.46247	0.41888	**0.56066**
	lkml-reply	0.62362	0.61391	0.59975	0.63348	0.64124	0.60591	0.59275	0.61802	0.64213	0.62009	0.55339	0.55099	0.61258	0.59637	0.61284	0.58421	**0.67979**
	mit	0.76332	0.78212	0.73823	0.80679	0.85671	0.85318	0.82605	0.84933	0.78789	0.80367	0.70919	0.71195	0.78432	0.82272	0.79426	0.78081	**0.88564**
	radoslaw-email	0.76087	0.77557	0.77527	0.73327	0.83567	0.84342	0.74869	0.83692	0.77204	0.78579	0.68831	0.69397	0.77181	0.78296	0.7745	0.77123	**0.85601**
	Eu-core	0.83703	0.85907	0.84222	0.67823	0.94347	0.9366	0.68971	0.91513	0.83978	0.88699	0.57749	0.57935	0.86718	0.85292	0.86294	0.84557	**0.94618**
BAL ACC SCORE	fb-forum	0.54835	0.5429	0.56522	0.66052	**0.89896**	0.88562	0.72192	0.883	0.80712	0.86608	0.58827	0.58486	0.57316	0.57952	0.58	0.5437	**0.89724**
	CollegeMsg	0.52372	0.51931	0.52908	0.58744	0.64509	0.6361	0.59461	0.6343	0.62059	0.65567	0.56756	0.55507	0.51515	0.54262	0.52903	0.51796	**0.68935**
	mathoverflow	0.69648	0.69289	0.69401	0.7129	0.71073	0.69656	0.69398	0.70433	0.71364	0.70921	0.70367	0.70235	0.68645	0.68878	0.68306	0.65557	**0.72556**
	lkml-reply	0.77974	0.77667	0.77933	0.78037	0.7836	0.77446	0.7668	0.77557	0.78426	0.77799	0.76974	0.76986	0.77608	0.77064	0.77115	0.75823	**0.79538**
	mit	0.8239	0.83665	0.80935	0.87627	0.89485	0.89556	0.89519	0.90508	0.86094	0.88141	0.77043	0.77794	0.84256	0.88082	0.85917	0.84149	**0.93635**
	radoslaw-email	0.82358	0.85081	0.85139	0.80656	0.91189	0.91729	0.82845	0.91648	0.84163	0.85484	0.75666	0.75754	0.84853	0.85311	0.85151	0.83904	**0.93462**
	Eu-core	0.88986	0.92031	0.90955	0.73316	0.97597	0.97245	0.7588	0.96021	0.90937	0.94579	0.60775	0.61027	0.92369	0.91676	0.92172	0.89413	**0.98174**
AUC	fb-forum	0.56679	0.53809	0.56503	0.71631	0.93072	0.91211	0.78449	0.92463	0.86116	0.91647	0.62157	0.61351	0.57034	0.56429	0.56219	0.54067	**0.93679**
	CollegeMsg	0.5315	0.53311	0.53833	0.63617	0.72636	0.66714	0.64998	0.69883	0.66038	0.70837	0.6071	0.59588	0.52584	0.54575	0.54028	0.50764	**0.77693**
	mathoverflow	0.70251	0.69513	0.69605	0.73234	0.73484	0.69664	0.70046	0.721	0.72883	0.71613	0.71903	0.71704	0.68831	0.68661	0.68775	0.65683	**0.75081**
	lkml-reply	0.79263	0.78688	0.7878	0.80068	0.81174	0.78468	0.77877	0.79931	0.80387	0.78647	0.79009	0.79	0.78591	0.77431	0.78234	0.76434	**0.81885**

TABLE 4.3: Performance of PWAF model and its machine learning variations - Linear Discriminant Analysis (LDA)

	Dataset	CN	AA	JC	PA	SP	COSP	ACT	MFI	LOCALP	L3	SRW	LRW	CCLP	CCLP2	NLC	CARCN	PWAF-LDA
	mit	0.53799	0.5371	0.4158	0.62971	0.47188	0.65641	0.5187	0.31194	0.60117	0.6747	0.35685	0.3553	0.55254	0.64853	0.53299	0.56349	**0.75327**
	radoslaw-email	0.57238	0.5677	0.55409	0.51833	0.48064	0.48561	0.36279	0.47914	0.58258	0.61289	0.29318	0.2896	0.58691	0.57871	0.59229	0.55809	**0.79059**
	Eu-core	0.69802	0.72893	0.69267	0.44424	0.80637	0.69356	0.43426	0.65798	0.71609	0.76965	0.18561	0.1829	0.77852	0.74602	0.77437	0.67771	**0.85809**
AUPR	fb-forum	0.23946	0.25541	0.20252	0.42755	0.34255	0.72993	0.26257	0.68498	0.51841	0.8041	0.26587	0.26556	0.25549	0.31112	0.31613	0.27514	**0.84609**
	CollegeMsg	0.22371	0.22471	0.18597	0.32005	0.19907	0.45868	0.28428	0.40066	0.35068	0.47724	0.2755	0.2776	0.23308	0.26526	0.24147	0.27878	**0.52246**
	mathoverflow	0.62735	**0.62794**	0.58091	0.58111	0.3123	0.59938	0.27707	0.46612	0.60031	0.6068	0.42895	0.42791	0.60936	0.62549	0.60691	0.62079	0.60175
	lkml-reply	0.71532	0.71801	0.6783	0.68944	0.39848	0.69963	0.35767	0.61381	0.70005	0.70486	0.5389	0.53902	0.72862	0.71826	**0.731**	0.71683	0.7139
	mit	0.61682	0.60343	0.5158	0.67197	0.70853	0.71921	0.5244	0.36725	0.64324	0.69188	0.36316	0.37142	0.59368	0.68271	0.57133	0.59792	**0.74881**
	radoslaw-email	0.59416	0.59109	0.59477	0.54504	0.6124	0.60397	0.44715	0.52932	0.5804	0.58984	0.40827	0.40818	0.60426	0.57634	0.58832	0.53681	**0.73633**
	Eu-core	0.67228	0.69309	0.70004	0.44445	0.81615	0.73356	0.37672	0.66485	0.67959	0.72763	0.29307	0.28397	0.71894	0.69017	0.68286	0.58039	**0.84755**
F1 score	fb-forum	0.25379	0.28218	0.28825	0.43503	0.56224	0.73563	0.38556	0.68102	0.51958	0.77557	0.30733	0.30323	0.24965	0.30416	0.27839	0.16996	**0.81074**
	CollegeMsg	0.19571	0.19457	0.26586	0.3416	0.28697	0.423	0.29501	0.37294	0.36441	0.4609	0.28314	0.28435	0.14112	0.2384	0.13232	0.09139	**0.4874**
	mathoverflow	0.51425	0.50784	0.50026	0.47322	0.42873	0.49188	0.27491	0.42103	0.50594	0.4949	0.43583	0.43453	0.34051	0.43936	0.25202	0.25007	**0.53207**
	lkml-reply	0.60154	0.60014	0.59373	0.54882	0.46806	0.65282	0.34233	0.55885	0.59223	0.58465	0.49324	0.49593	0.49192	0.56001	0.41488	0.32993	**0.66413**
	mit	0.5474	0.54588	0.42291	0.63542	0.5506	0.66118	0.52484	0.3186	0.60931	0.67883	0.36197	0.36007	0.5605	0.65508	0.54182	0.57129	**0.7574**
	radoslaw-email	0.57486	0.57013	0.55614	0.52088	0.63962	0.48667	0.36424	0.48146	0.58506	0.61455	0.29418	0.29057	0.58939	0.58091	0.59466	0.56053	**0.79144**
	Eu-core	0.69973	0.73049	0.69538	0.44646	0.85283	0.69599	0.43743	0.66041	0.71772	0.77092	0.18778	0.18514	0.77961	0.74692	0.77519	0.67616	**0.85987**
AVG PRECISION	fb-forum	0.23264	0.25103	0.19537	0.43066	0.34788	0.72767	0.25517	0.68792	0.52203	0.8051	0.26971	0.26941	0.24749	0.30287	0.30558	0.21679	**0.84638**
	CollegeMsg	0.20829	0.20854	0.1735	0.32323	0.18333	0.43415	0.17851	0.39868	0.34897	0.47535	0.27753	0.27993	0.21926	0.25205	0.22601	0.17927	**0.49601**
	mathoverflow	0.48557	0.49309	0.44801	0.53519	0.27174	0.49813	0.23527	0.41937	0.53933	0.54626	0.38592	0.38485	0.47083	0.48971	0.46751	0.42203	**0.54442**
	lkml-reply	0.60754	0.61137	0.58481	0.64078	0.36184	0.64621	0.31182	0.5684	0.64118	0.64582	0.49518	0.49501	0.63292	0.61555	0.63654	0.57877	**0.66515**
	mit	0.76514	0.76202	0.69843	0.80204	0.84234	0.83697	0.72352	0.58585	0.78194	0.81029	0.59639	0.59988	0.75838	0.80072	0.74676	0.75203	**0.86322**
	radoslaw-email	0.76788	0.76681	0.78123	0.73274	0.7836	0.7693	0.6846	0.73092	0.76052	0.75897	0.6353	0.63556	0.76989	0.747	0.75116	0.70357	**0.84165**
	Eu-core	0.81257	0.81941	0.83384	0.67412	0.93053	0.82794	0.65849	0.79604	0.81151	0.84062	0.53079	0.53016	0.81678	0.80517	0.78425	0.72236	**0.92257**
BAL ACC SCORE	fb-forum	0.54533	0.55968	0.56113	0.66877	0.78843	0.81631	0.65423	0.78179	0.72541	0.84451	0.58047	0.57808	0.55307	0.58406	0.57486	0.53674	**0.87183**
	CollegeMsg	0.53117	0.5302	0.49791	0.60804	0.53844	0.6422	0.54315	0.61905	0.61796	0.66275	0.56045	0.56396	0.52201	0.55643	0.52749	0.50969	**0.67919**
	mathoverflow	0.68072	0.6763	0.67866	0.65805	0.67197	0.66799	0.5677	0.6538	0.67414	0.66765	0.65502	0.65441	0.60266	0.64219	0.57183	0.57127	**0.71417**
	lkml-reply	0.72227	0.72087	0.73151	0.6927	0.69234	0.76986	0.60444	0.73831	0.71625	0.71091	0.68249	0.68375	0.66424	0.69759	0.63117	0.59878	**0.7844**
	mit	0.8268	0.82563	0.7508	0.86383	0.83425	0.86587	0.80906	0.63673	0.85404	0.88133	0.65016	0.65111	0.82802	0.87661	0.82166	0.81008	**0.92137**
	radoslaw-email	0.8566	0.85943	0.85467	0.8238	0.86816	0.74247	0.77618	0.80672	0.85529	0.86227	0.70919	0.70523	0.86305	0.85122	0.85993	0.85565	**0.92102**
	Eu-core	0.92041	0.92629	0.92195	0.75647	0.97208	0.82787	0.76907	0.85924	0.92553	0.94488	0.54225	0.54483	0.94184	0.92948	0.93806	0.88557	**0.97428**
AUC	fb-forum	0.55702	0.57263	0.56117	0.73531	0.79898	0.82987	0.6893	0.82912	0.81781	0.91001	0.64179	0.63892	0.58098	0.61516	0.59876	0.53911	**0.93838**
	CollegeMsg	0.54155	0.54218	0.50968	0.6275	0.5426	0.65494	0.53291	0.62624	0.64627	0.67248	0.5928	0.59353	0.53319	0.5569	0.52797	0.50044	**0.75461**
	mathoverflow	0.6969	0.70145	0.70104	0.71988	0.68202	0.70586	0.66913	0.69228	0.72636	0.72726	0.70709	0.70616	0.69687	0.70056	0.69262	0.65847	**0.7378**
	lkml-reply	0.76008	0.76053	0.78945	0.7825	0.76296	0.79035	0.72285	0.78774	0.77976	0.7796	0.78669	0.7865	0.78943	0.76256	0.79255	0.75058	**0.81178**

well-known dynamic networks with five evaluation metrics, the experimental results showed that PWAF machine learning variations yield higher accuracy.

4.3.1 Performance of PWAF model and its machine learning variations- XGBoost (XGB)

In this section, we assess the performance of the PWAF model and its variations on feature sets using XGB as a training and testing model. We have used 50 estimators with learning rate of 0.01 as settings for this classification.

Table 4.1 compares the performance of the PWAF-XGB model against that of other similarity indexes using XGB. In terms of AUPR, COSP and SP gives better performance on mit, radoslaw-email and Eu-core daases. PWAF-XGB gives better result on fb-forum dataset among all other methods. SP gives better performance on CollegeMsg dataset. PWAF-XGB gives better performance on mathoverflow and lkml-reply datasets. When the F1 score is included, PWAF-XGB outperforms all other techniques. On Eu-core dataset, PWAF-XGB gives similar result as SP. In terms of AVG PRECISION, COSP gives better performance on mit dataset whereas SP gives better performance on radoslaw-email and Eu-core dataset. On the fb-forum CollegeMsg, mathoverflow, and lkml-reply datasets, PWAF-XGB outperforms all other approaches by a significant margin. The best performing approach among all algorithms in terms of BAL ACC SCORE is PWAF-XGB on all datasets. SP and COSP also produce comparable results. JC gives better performance on lkml-reply datasets. On the mit dataset, PWAF-XGB, SP, and COSP perform similarly to and better than other approaches in terms of AUC. PWAF-XGB gives better performance on all other datasets.On lkml-reply datasets, JC, SP, and PWAF-XGB perform comparably to other approaches and produce superior results.

4.3.2 Performance of PWAF model and its machine learning variations- Random Forest Classifier (RFC)

In this section, we evaluate the performance of the proposed PWAF approach. We have used 100 estimators as setting to create this classifier and used default Scikit-Learn [156] implementation.

Using RFC, Table 4.2 compares the PWAF-RFC model's performance to that of various other similarity indexes. On all datasets tested, PWAF-RFC beats all other approaches when the AUPR score is taken into account. On the mit and radoslaw-email datasets, PWAF-RFC outperforms all other algorithms in terms of F1 score. On the Eu-core and fb-forum datasets, SP outperforms all other algorithms. On the CollegeMsg and mathoverflow datasets, PWAF-RFC produces better results. On the lkml-reply dataset, PWAF-RFC, AA, JC, SP, CCLP, and NLC produce similar results. On all datasets, PWAF-RFC surpasses all other approaches in terms of AVG PRECISION. On all datasets except fb-forum, where it performs similarly to SP in terms of BAL ACC SCORE, PWAF-RFC outperforms all other approaches in terms of BAL ACC SCORE. On all datasets, PWAF-RFC surpasses all other algorithms by a considerable margin when it comes to AUC.

4.3.3 Performance of PWAF model and its machine learning variations- Linear Discriminant Analysis (LDA)

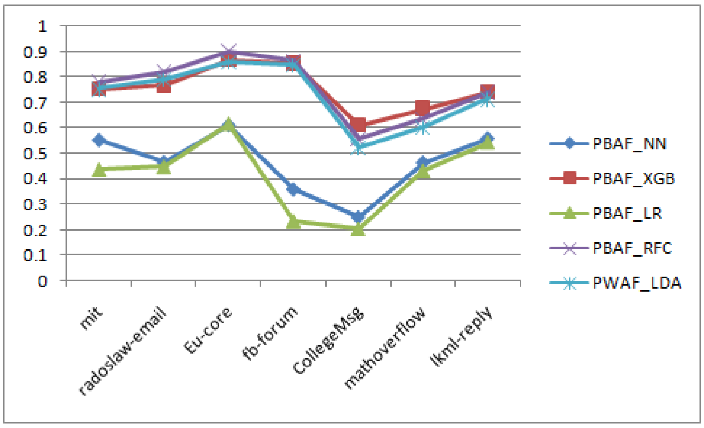
In this section, we'll see how well the PWAF approach and its variations perform on feature sets using the LDA training and testing model. We have used default Scikit-Learn [156] implementation for this classifier.

Table 4.3 compares the performance of the PWAF-LDA model to specific features in the LDA machine learning classifier. On the mit, radoslaw-email, Eu-core, fb-forum, and CollegeMsg datasets, PWAF-LDA performs better in terms of AUPR. AA and NLC outperform on mathoverflow and lkml-reply, respectively. On the mit, radoslaw-email, Eu-core, fb-forum, CollegeMsg, mathoverflow, and lkml-reply datasets, PWAF-LDA performs better in terms of F1 score. On the mit, radoslaw-email, Eu-core, fb-forum, CollegeMsg, mathoverflow, and lkml-reply datasets, PWAF-LDA shows the best performance of AVG PRECISION. On the mit, radoslaw-email, Eu-core, fb-forum, CollegeMsg, mathoverflow, and lkml-reply datasets, PWAF-LDA achieves a higher BAL ACC SCORE. On the mit, radoslaw-email, Eu-core, fb-forum, CollegeMsg, mathoverflow, and lkml-reply datasets, PWAF-LDA performs better in terms of AUC.

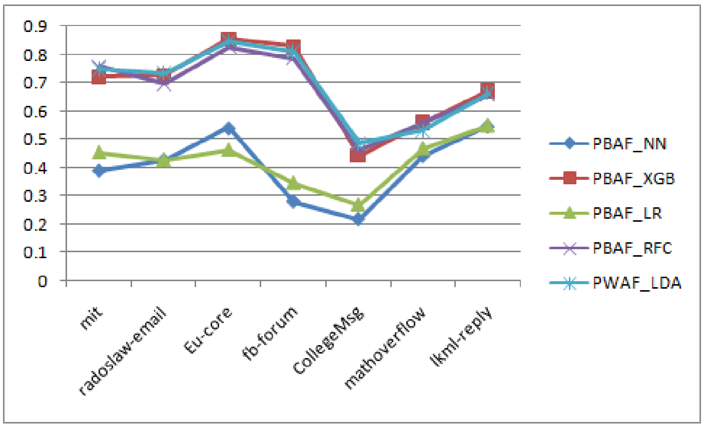
4.3.4 Performance comparison of PWAF model machine learning variations

In this subsection, we'll look at how different PWAF model machine learning variants performed on seven well-known dynamic networks using five different performance metrics. PWAF-NN, PWAF-LR, PWAF-XGB, PWAF-RFC, and PWAF-LDA are the machine learning variations employed. The comparison and analysis of several PWAF machine learning variations is shown in Figure 4.6. On the mit, radoslaw-email, Eu-core, and fb-forum datasets, PWAF-RFC outperforms PWAF-RFC in terms of AUPR. On the CollegeMsg, mathoverflow, and lkml-reply datasets, PWAF-XGB produces the best results. On the mit, radoslaw-email, Eu-core, and fb-forum datasets, PWAF-RFC outperforms all other variations in terms of AUPR. On the CollegeMsg, mathoverflow, and lkml-reply datasets, PWAF-XGB produces the best results. On the mit and radoslaw-email datasets, PWAF-RFC and PWAF-LDA produce the best on F1 score. On the Eu-core and fb-forum datasets, PWAF-XGB performs best. On the CollegeMsg, mathoverflow, and lkml-reply datasets, PWAF-LDA, PWAF-RFC, and PWAF-XGB yield the greatest results. On the mit, radoslaw-email, Eu-core, fb-forum, CollegeMsg, mathoverflow, and lkml-reply datasets, PWAF-RFC outperforms all other variations in terms of AVG PRECISION, BAL ACC SCORE, and AUC.

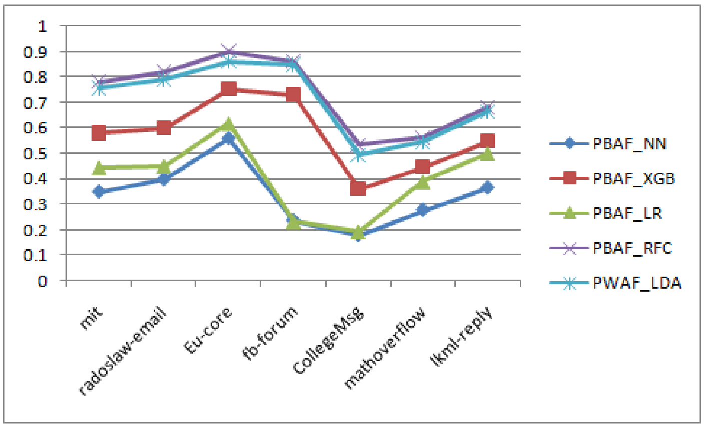
FIGURE 4.6: Performance Comparison among PWAF machine-learning variations



(A) AUPR

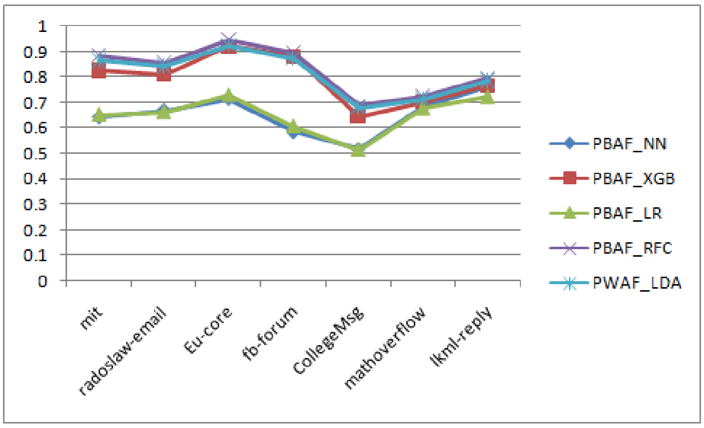


(B) F1 score

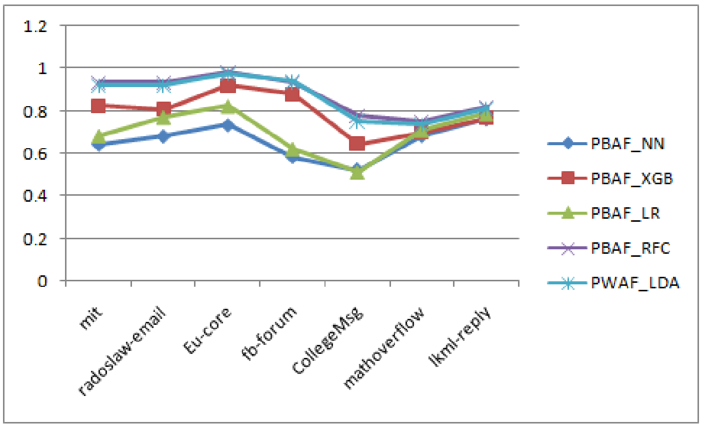


(C) Average Precision

FIGURE 4.6: Performance Comparison among PWAF machine-learning variations (contd..)



(D) Balance Score



(E) AUC

4.3.5 Comparison of PWAF variations with state-of-the-art methods

In this section we compare the performance of proposed Path Weight Aggregation Feature (PWAF) for Link Prediction in Dynamic Networks with five state-of-the-art approaches. The result of five state-of-the-art methods are compared to the performance of our proposed optimal machine learning variations, namely PWAF-XGB, PWAF-RFC, and PWAF-LDA, in Table 4.4. These results are compared on seven well-known open dynamic datasets using five performance evaluation metrics: AUPR, F1 Score, AVG PRECISION, BAL ACC SCORE and AUC. PWAF-RFC can be seen to be overall the best performing algorithm on all datasets. In terms of the AUPR evaluation measure,

TABLE 4.4: Performance comparison of PWAF machine learning variation with state-of-the-art methods

	Dataset	RA-RFC	PROXM	Node2Vec	WEAK	LGQ	PWAF-XGB	PWAF-RFC	PWAF-LDA
AUPR	mit	0.58164	0.37668	0.34857	0.55145	0.74443	0.75111	**0.78**	0.75327
	radoslaw-email	0.58913	0.43212	0.41985	0.33067	0.75981	0.76699	**0.81991**	0.79059
	Eu-core	0.74903	0.66827	0.7356	0.48863	0.86477	0.86499	**0.90094**	0.85809
	fb-forum	0.20723	0.58368	0.65555	0.38666	0.84059	0.85159	**0.86375**	0.84609
	CollegeMsg	0.22062	0.25977	0.42017	0.29695	0.609	**0.61114**	0.55706	0.52246
	mathoverflow	0.61734	0.56755	0.49082	0.54906	**0.67385**	**0.67461**	0.63686	0.60175
	lkml-reply	0.7181	0.72403	0.64506	0.61187	0.73993	**0.74067**	0.7399	0.7139
F1 score	mit	0.61092	0.23581	0.42454	0.41404	0.71414	0.72016	**0.75758**	0.74881
	radoslaw-email	0.57081	0.25144	0.42119	0.28952	0.71774	0.72509	0.69715	**0.73633**
	Eu-core	0.68485	0.48031	0.69285	0.43604	**0.85376**	**0.85368**	0.82693	0.84755
	fb-forum	0.25453	0.47254	0.57337	0.35492	0.81832	**0.82989**	0.78619	0.81074
	CollegeMsg	0.21895	0.24988	0.38145	0.29882	0.44578	0.44215	0.46203	**0.4874**
	mathoverflow	0.52522	0.49236	0.47727	0.54243	**0.55416**	**0.55717**	**0.55742**	0.53207
	lkml-reply	0.65232	0.56876	0.59426	0.63416	**0.67244**	**0.67096**	0.65899	0.66413
AVG PRECISION	mit	0.58341	0.38432	0.3569	0.38221	0.56731	0.57971	**0.78113**	0.7574
	radoslaw-email	0.5866	0.43367	0.42233	0.2934	0.59009	0.59996	**0.81697**	0.79144
	Eu-core	0.74572	0.63934	0.73792	0.37745	0.74961	0.75083	**0.89868**	0.85987
	fb-forum	0.19778	0.53808	0.65866	0.34988	0.71055	0.72856	**0.85941**	0.84638
	CollegeMsg	0.19928	0.19046	0.41533	0.25821	0.3669	0.36126	**0.53456**	0.49601
	mathoverflow	0.47005	0.48114	0.44387	0.49389	0.44408	0.44589	**0.56066**	0.54442
	lkml-reply	0.61696	0.69856	0.59778	0.57838	0.54653	0.54658	**0.67979**	0.66515
BAL ACC SCORE	mit	0.78267	0.78816	0.63489	0.65212	0.82818	0.82597	**0.88564**	0.86322
	radoslaw-email	0.77531	0.82508	0.6507	0.59196	0.80618	0.8101	**0.85601**	0.84165
	Eu-core	0.87033	0.93282	0.88257	0.68926	0.92357	0.92083	**0.94618**	0.92257
	fb-forum	0.54646	0.82043	0.78762	0.61863	0.87505	0.88229	**0.89724**	0.87183
	CollegeMsg	0.53498	0.64007	0.63855	0.56032	0.64732	0.64581	**0.68935**	0.67919
	mathoverflow	0.69241	**0.78381**	0.69452	0.71551	0.69913	0.70083	0.72556	0.71417
	lkml-reply	0.77711	**0.88832**	0.77343	0.78094	0.76838	0.76681	0.79538	0.7844
AUC	mit	0.83885	0.89392	0.68906	0.65496	0.82818	0.82597	**0.93635**	0.92137
	radoslaw-email	0.84999	0.91898	0.73568	0.60541	0.80618	0.8101	**0.93462**	0.92102
	Eu-core	0.9242	0.9552	0.94494	0.70495	0.92357	0.92083	**0.98174**	0.97428
	fb-forum	0.53959	0.79522	0.85596	0.64254	0.87505	0.88229	**0.93679**	**0.93838**
	CollegeMsg	0.54697	0.63843	0.71109	0.57384	0.64732	0.64581	**0.77693**	**0.75461**
	mathoverflow	0.69462	**0.78762**	0.71539	0.73463	0.69913	0.70083	0.75081	0.7378
	lkml-reply	0.78708	**0.92316**	0.80217	0.79669	0.76838	0.76681	0.81885	0.81178

PWAF-RFC outperforms all other algorithms on the mit, radoslaw-email, Eu-core, fb-forum, and CollegeMsg datasets, but PWAF-XGB outperforms all other algorithms on mathoverflow and lkml-reply datasets. When it comes to F1 scores, PWAF-RFC outperforms all other algorithms on the mit dataset, whereas PWAF-LDA outperforms all other algorithms on the radoslaw-email and CollegeMsg datasets. PWAF-XGB outperforms the competition on the Eu-core and fb-forum datasets. On the mathoverflow dataset, PWAF-RFC, PWAF-XGB, and LGQ produce similar results. On the lkml-reply dataset, LGQ and PWAF-XGB both perform well. In terms of AVG PRECISION, PWAF-RFC outperforms all other algorithms on the mit, radoslaw-email, Eu-core, fb-forum, and CollegeMsg datasets, as well as mathoverflow datasets, and PROXM outperforms all other techniques on the lkml-reply dataset. On the mit, radoslaw-email, Eu-core, fb-forum, and CollegeMsg datasets, PWAF-RFC outperforms every other approach in terms of BAL ACC SCORE. PROXM is a good performer on the mathoverflow and lkml-reply datasets. On the mit, radoslaw-email, Eu-core, and CollegeMsg datasets, PWAF-RFC outperforms all other approaches in terms of AUC, whereas PWAF-LDA outperforms all other methods on the fb-forum dataset. On mathoverflow and lkml-reply datasets, PROXM provides a superior result. Based on these results, we can infer that of all the machine learning classifiers we've tested, PWAF-RFC is the best performing variation.

4.4 Conclusion

We attempt to solve the link prediction problem in dynamic networks using an enlarged feature set which represents different levels of node information in this research. The Path Weight-Based Aggregation Feature (PWAF) is a new feature that we propose. In addition to the recommended Path Weight-Based Aggregation Feature (PWAF), several topological properties of the networks (Local, Global, and Quasi-local), as well as Clustering Coefficient based features are taken into consideration for feature generation. The Level-2 node clustering coefficient (CCLP2) is one of the features used to improve prediction. For link prediction, many machine learning models are used to make predictions using this rich feature set, including Neural Network (NN), Logistic Regression (LR), XGBoost (XGB), Random Forest Classifier (RFC), and Linear Discriminant Analysis (LDA). The experiments are carried out on seven different well-known dynamic networks data sets in terms of five performance evaluation metrics, including AUPR, F1-score, AVG PRECISION, BAL ACC SCORE, and AUC, and the results show that our proposed method and its variants outperform state-of-the-art methods. Among all algorithms and state-of-the-art approaches, PWAF-RFC is the top performer. In addition, PWAF-XGB also provides superior performance among individual features as well as state-of-the-art methods.