Social Network Analysis
Prof. Tanmoy Chakraborty
Department of Computer Science and Engineering
Indraprastha Institute of Information Technology, Delhi

Lecture - 73
Concluding Panel Discussion

Hello everyone, welcome you all in the for the last time to this course on Social Network Analysis. So, you may wonder what is the agenda of today's discussion I have already concluded the course in the last lecture. So, and we have already discussed the you know need of social network analysis in general, we have discussed the classical part of social network analysis including say growth models and the old school statistics etcetera.

We have also discussed modern social network analysis you know using deep learning to understand graphs in a better way right, but you know whatever we have discussed so far. So, you have heard you have listened to me from a instructors point of view right you may not have ideas about what a student thinks about this course, of course, you have your own opinion in mind about this course.

But I thought you know this would be an interesting idea to have a small you know discussion round with you know a set of students right who are actually working on social network analysis and may have better ideas you know than me about this course about this topic in general.

So, you know today I mean we are you know going to have a very small you know kind of candid discussion right, this is kind of a you know coffee with Karan types session. So, it is not like that kind of thing, but we wanted to have it very candid so that you also understand you know from students point of view why this is an interesting topic this is an interesting you know academic course this is an interesting research topic right.

So, it is my pleasure to introduce you know introduce to you today's panellists I would say I mean they are all my students, but for today's discussion they are all panellists. So, I have with me three of my students. So, they are PhD students working in my lab. So, Shivani you know Shivani is the TA of this course she is also one of the senior you know members of this of my lab.

We have Karan, Karan is I would say semi senior kind of folk in the lab he also works in graph. So, he will talk about his work. Sara is another senior student in the lab, she is also a prime minister fellow she works in social network I mean in the intersection of social network and graphs in general. So, and this is going to be a very frank you know very kind of candid discussion ok.

So, let me introduce today's panellists although they are my PhD students, but for you know today's discussion they are panellists. So, you know Shivani, Shivani is the TA of this course, she is also the I mean one of the senior members in the lab, she works in the intersection of natural language processing and graphs.

Karan is another PhD student working in the lab, he is kind of a you know the semi senior I would say PhD student, he works in graphs particularly deep learning for graphs. And Sara, Sara is an another senior PhD students working in the lab and she is also a prime minister fellow she works in the intersection of social network analysis and graph mining in general.

So, this is going to be a very you know frank candid discussion just to understand students perspective about the course and why this is an interesting important research area right. Although, I you know kind of motivated you a lot that if you learn this course you would be able to you know to be interviewed by companies like Google, LinkedIn right top notch companies working in social network Facebook, Meta, Twitter and so on, but let us hear it from the students right what they think about it.

So, let me start with Karan right. So, Karan, what do you think about this course? Why you think that you know social network analysis is a graph mining per se right is an in interesting topic in today's arena particularly in today's arena, where I mean we see a huge amount of data available publicly right. We can scrape data from Facebook; we can scrape data from Twitter, of course, with some restrictions right. Why do we think that this is important and people should pursue this as one of the career options?

Karan: Yeah. So, this course specifically regarding social network analysis is very important let us say you have a lot of data right that is from Twitter. So, there are many activities going on there and for example, this hate speech and fake news, these are very prevalent terms in today's world. So, we do want to address these challenges, for that we need a network analysis right and that has some properties and some algorithm that have been discussed in this course.

So, we do want to put our focus on these specific nuances and details. So, that we can address these challenges and for example, right you discussed about the network properties like degree centrality, articulation points and similarly if I talk about a large level perspective let us talk about community detection, people from eco chambers also there.

So, we do want to identify these things. So, for that we need to learn these techniques first. So, I mean that is very good. So, but can you tell us very briefly about your research, what is the research area that you are currently working in?

Karan: Ok. So, I work on algorithmic side of graph that is graph representation learning. So, in that our aim is to learn embeddings that can represent nodes in a graph in a vector space. So, essentially what we do? We conduct amalgamation of deep learning and graphs, so that we can devise new algorithms that can operate a specifically on graph structure data. So, to perform successfully and task like node prediction, node classification, edge prediction, graph classification etcetera may become reductions and so on. So, I work generally in this domain yeah.

Ok.

Shivani: Yeah. So, basically that you have mentioned so, this graph representation learning is it like similar to normal representation learning in general.

Karan: Ok, yeah it has a same notion of achieving and doing the things, but if you see graphs are very difficult and exciting to handle than other type of data.

Shivani: Right.

Karan: Let us say graphs are not sequential in nature right.

The other type of data you have let us say we have frames in a video. Maybe you have pixels in an image or sentences in a paragraph or words in a sentence. So, graphs are not sequential there. So, they have very specific property of isomorphism. That are only specific to graphs.

Shivani: Ok.

Karan: And also the one thing that is very specific to graph is that you can increase the complexity of the data enormously just by adding the number of edges between already existing data points. That is nodes. You keep them fixed.

Shivani: Ok.

Karan: So, that is not the case you observe with images videos or texts.

So, you know just to just to interrupt here, audience if you remember I exactly mentioned this point in the graph representation learning chapter that why you know graph representation learning is different from say language representation learning right, what is the challenge in graphs right.

We discussed that you know unlike say a speech right or a text where you have a sequence sequential data available right, you know that ok, the word say I am a boy the word am is at the second position right in the particular sentence in the graph there is no such notion of position right of a node in the graph right.

So, and Karan rightly mentioned and this is one of the emerging areas that is going there in the research community right, from the static graph, temporal graph, heterogeneous graph, right attributed graph you have plenty of information's, how do you in inject all this information, how do you incorporate them together to come up with better representation for the downstream task.

Now, the downstream task can be you know it can the representation can be task dependent the representations can also be task independent right and as Karan mentioned I think one of the major focus of his research is to learn the representation right in an appropriate manner particularly in the temporal setting ok.

Karan: Yeah. So, you have a question.

Sara: Yeah. So, you talked about graph as a data structure being unique, then how do you incorporate these unique properties and network metrics into graph designing your graph representation learning algorithms?

Karan: Ok. So, you essentially trying to ask that ok how are these network properties or measures important in designing the graph.

Sara: Yes.

Karan: Representation.

Sara: Definitely.

Karan: Ok. So, if you see right most of the graph representation learning algorithm that have been developed in the past few years they focus mainly on neighbourhood aggregation the base around that. So, unlike right deep learning like architectures or models whatever you say. So, these are not like here we cannot go on keep on increasing the layers.

Because in the end, we will end up learning nearly the same representation for every node in the graph. So, it is how it is different. So, here we needed to focus on this thing. So, now, what will happen is we will and same representation, now a network property is responsible for this and that property is small world property. So, if we already know that ok small world property you can reach every node in the graph from any particular node in 4 to 6 steps. So, you would not be right increasing a number of likes too much.

Sara: No much.

Karan: And then just see ok my algorithm is performing poor. So, that is why it is important. I will give you another example from a real world data set that we worked on. So, we have a data set where the and task was to do node classification. So, and we ran standard algorithms like graph convolutional network, graphs edge and graph attention networks and they perform miserably and we were surprised. So, then we dug up the issue and we found that there was no homophily in the network.

So, yeah we should care about the different network properties and measures and other techniques while designing and also utilizing these algorithms.

Karan: And I think they have been discussed in this course very well.

Sara: Ok.

So, again I hope you have you have understood some of the terminologies if not the entire discussion, but at least you have you know the you have brightly pointed out the terms like homophily and all these stuffs right that we have discussed in the course right, thanks Karan. So, let me now move to Sara. So, Sara tell us briefly about your research your research interests your research work in general.

Sara: Sure. So, if we look at the world around us as we were discussing earlier it can be mapped in terms of various entities that are interacting with each other and these interactions are very dynamic in nature.

Similar thing happens when you look at the online world where you have various heterogeneous entities like you have users you have products, you have content and all of these entities are interacting with each other and my area of research which is called social computing aims to understanding these various interactions and you know trying to on earth or properties of these interaction as well as what can be done once we understand these interactions.

So, my area of research which is social computing within that I work specifically in the area of analyzing the networks for hate speech detection as well as diffusion.

Karan: Ok. So, I have one question.

Sara: Yeah.

Karan: So, I am just trying to relate my work with yours. So, I think we can use right graph representation learning techniques maybe let us say you mentioned hate speech, to identify hate speech users or fraudulent users from some benign ones can we use that.

Sara: Yes definitely. So, see once we are able to map these user interaction networks in form of a graph then we can run all sort of graph representation learning algorithms on top of that and the embeddings that we obtained from this then we can you know use it for various downstream task as you were mentioning one of them would be like classifying a user as benign or harmful or maybe you want to detect whether a user is a bot user or a genuine user.

And not only that not only like specific to just classifying the network embedding you can we have personally observed in our research that when you combine these network embedding along with let us say the textual embeddings of a post, you can improve the performance of hate speech detection algorithms and this goes back to the property of homophily where we know that people who are more likely to interact with other hateful users are likely to also post hateful content.

Shivani: So, all the task that you just mentioned they are essentially you know classification or detection tasks right, but can we use this user interaction networks to go beyond classification and perform like more complex tasks maybe in the network?

Sara: Yes, definitely, we can do that like at the end of the day everything boils down to classifying a user or a content into some category, but networks are useful for not just categorizing specific groups or of users or entities, we have studied in this course specifically that networks are also useful for understanding how the information is spreading among these various entities. So, the concepts of cascade and information diffusion can be studied by understanding these user interactions.

Sara: And take the example of fake news we know that it goes viral very easily. So, if you have understood, ok these are the central nodes in my network and you know that they are more likely to spread fake news then you can develop your intervention strategies in a manner that you know prevents the spread of this fake news. So, understanding diffusion through networks is another area of research.

Shivani: Correct.

So, you know thanks Sara. So, I mean if you if we summarize the discussion of Sara's research. So, Sara mostly focuses on understanding a particular offensive content which is hate speech and how we detect it, how we stop the spread of hate speech, can we come up with you know different intervention strategies to combat you know online hate speech in general right.

Now, let me move to Shivani, Shivani's work is very interesting she works in the intersection of NLP and I mean mostly NLP, but her research also you know needs a lots of technique from graphs right. Now, again I motivated you in in one of the chapters that you know the graph you know by itself, I mean if you think of graph I mean one can assume that this is basically a theoretical you know domain right.

It has less practical application, but you know it has a lot of places where you where you can use graphs in different applications, one of these applications is language technology natural language processing. So, let me you know ask Shivani. So, can you briefly you know tell us about your research and why you feel that you felt that graph could be one of the I mean one of the components of your research?

Shivani: So, basically my research domain is about you know analyzing the discourse between different speakers right. So, I work in the area of these dialogue analysis and identifying and you know explaining the different aspects of the dialogue like emotions and sarcasm and profiling different speakers right. So, when we talk about graphs so, basically graphs requires some kind of nodes and some kind of relationship between the nodes to for the edges to form right.

So, in the terms of dialogues so what we have as an input is basically a sequence of utterances which are uttered by various different speakers. So, now, these speakers and these utterances they basically exist with some kind of relationship among each others. So, the speakers might have some different kind of relationship among others and the utterances can have some other kind of relationship with each other.

For instance like a simple relationship can be a temporal relationship that is one utterance is occurring after the another entrance. So, we can you know exploit these relationships and can explain or can present these dialog input as the form of graph and then as we just discussed can use the existing graph representation methods or another other techniques that we have studied in this course to basically analyze these structure and then use it for the end task NLP task like sentiment classification and so on ok.

Sara: I had a question. So, you talked about the concept of utterances.

Shivani: Right.

Sara: But they where they have relationships, what about applying the concept of graphs in task that do not have these you know relationship explicitly present?

Shivani: Right ok yeah. So, I think what you are trying to ask is like can we apply graph structure to stand alone text like say movie reviews or tweets right. So, yeah so, definitely we can, because you know national language itself the words that are present in the national language they often occur with some kind of relationship with each other. So, these words in a sentence they can have like a syntactical relationship which can you know arise from the grammar of the language or they can have a semantic relationship for instance like the relationship between the POS tags of the different words in the sentence.

So, then in NLP we have different kinds of parser which basically you know takes as input the input sentence and then results in a network of this syntactic and semantic relations. So, the output of this parser or the graph or the network structure that we get from these parser then again this graph can be used as a separate feature or itself to again like perform the task like we just mentioned like sentiment classification or sarcasm detection and so on. So, yeah so we can use this graph structure for standalone text as well.

Karan: Ok. I have some question. So, I think I we can use graphs right as a source of external knowledge for NLP task because they can add a lot to the information pool already out there.

Shivani: Correct.

Karan: So, will you be able to share your views on that.

Shivani: Yeah sure. So, see I guess you are talking about knowledge graphs right.

Karan: Yeah.

Shivani: Yeah. So, the area of knowledge graphs and using those in NLP task it is basically an area that is currently being you know well explored in the NLP domain because you know let us just understand it intuitively. So, for example, if we are talking and you mentioned say like you admire the work of Leonardo da Vinci then me as like I should be aware like who Leonardo da Vinci is or maybe some of his works to understand your statement clearly right.

So, same goes for like say online dialogue agents right. So, if you are talking to a chat bot and you mentioned the same thing that you admire the work of Leonardo da Vinci then the chat bot must also you know should have some kind of background knowledge or some kind of context about who Leonardo da Vinci is or what are of what are his works.

So, yeah so, for incorporating this kind of knowledge for that knowledge graph such as you know concept net or for incorporating some kind of common sense knowledge so, concept graphs such as a commit they are basically they can be used for NLP task to incorporate these knowledge and can answer your you know the end task whatever you want to see.

Karan: So, they will essentially help us in lifelong learning right I think.

Shivani: Yeah basically. So, like we learn with our experience.

Karan: Correct.

Shivani: So, see we cannot give such a long time to the machines right. So, to just you know to provide the background knowledge beforehand to the machine. So, for that knowledge graphs can be helpful.

Karan: Ok fine.

Shivani: Yeah.

So, I mean if you have noticed, so she has mentioned a terminological knowledge graph, although I was I think I did not mentioned this term in the course, but we discussed a related terminology which was you know heterogeneous graph or you know something like that.

So, if you think of knowledge graph is basically I mean you can think of it as I mean you can I mean you can easily define it using a tuple you have. So, you have you know two objects and two objects are connected through a relation right and that object the these two objects right they can have their own properties this relation can also have its own property right.

For example say, there is a huge knowledge graph and there are two nodes two entities in the knowledge graph one is Barack Obama another is the US and the relation is you know was President right or this knowledge graph there is India as an entity and Narendra Modi is another entity and the relation is Prime Minister right and something like that.

So, you know you can think of this huge knowledge right in terms of graphs and then you can you know you I mean you can run whole bunch of the state of the art again representation learning techniques to understand to represent the relations to represent nodes and so on and so forth for different tasks.

So, let me ask a few more questions right. So, since you have particularly you know Shivani and Sara since you have already faced many interviews right what kind of questions are generally asked and what is the expectation right and I mean what kind of things people actually expect from a student to have at this moment about say let us say you are you are posing yourself as a graph learning research I mean researcher working graphs right. So, what kind of expertise people generally expect from a candidate from a student, what do you think Sara?

Sara: So, one of the interesting things about studying social network analysis is that it covers a wide range of topic. So, you get to learn about statistical methods. So, you know you have to improve your understanding of that, at the same time you have the other spectrum of graph representation learning where you have to have a background knowledge of deep learning systems, apart from that you know when we are analyzing graph we are using some libraries that we do not usually use.

For example in python we work extensively on network x. So, it gives you a wide range of what you can say knowledge and expertise in one particular course for example, in my interviews I have seen that when we talk about having a background knowledge in graph learning they expect us to be proficient with both deep learning as well as hands on in libraries like network x.

Yeah.

Sara: So, that is what I have observed that they expect us to know.

What do you think Shivani?

Shivani: Yeah. So, especially if you talk about the NLP domain and stuff so, basically with the boom of this social networks and all. So, many you know many companies or many interviewers they expect you to have some kind of a knowledge about the you know about how information spread throughout these networks. So, that you know if the interviewer they deal with these kinds of networks.

So, yeah they have that kind of expectation and as Sara correctly mentioned that the knowledge of statistical graph representation as well as this deep learning knowledge that how you can basically exploit this graph representation that we have to perform some kind of end task. So, that is some kind of knowledge that the interviewers they expect nowadays that we must have because essentially all the data can be you know expressed as a form of graph and then the methods can be used over that.

So, what I understood is that you know hands on experience is of course, needed, you should be an excellent coder like to be able to crack this interviews to be able to say deal with large scale networks, at the same time you should also have a decent knowledge you know about the theoretical understanding of all these techniques right.

So, guys be careful. So, whatever techniques I have discussed so far particularly in the network growth models and you know graph learning models you should be able to understand those techniques properly, if the video was not sufficient if you think that the that the video is not sufficient you just look at the original papers right.

For example, the node two week paper or the deep work paper or the graph GCN paper right to understand things thoroughly that is very important, do not think that you know just take something you know import a library run it right through some results and the reviewers will be the interviewers will be impressed with the results right of course, hands on experience is needed.

So, thanks guys, thanks a lot for the comments, I think this was really helpful to our audience particularly students and practitioners, who are still thinking about you know whether this is a really an interesting area to explore further and I also again like to thank all the listeners all my audience who have gone through the this long journey right of 12 weeks and hopefully you have understood most of the concepts that we have discussed.

As I mentioned I mean you I mean if you are really interested to work on this topic there are multiple publicly available sources from which you can get data you can get codes right just run it see the results, if you think that improvement is possible build your own model and try to improve it and if possible publish it in a good venue ok.

There are top venues in this area conferences like SIGKDD, world wide web, right ICDM, NeurIPS and triple AI these are top conferences where people publish their good work right, strongly encouraging you guys to focus on those venues read state of the art research and update your skills ok.

And hopefully you know we will meet some other time to discuss further about this particular area this kind of exciting area and you feel free to reach out to me with I mean using my email id and you know my other social media handles.

Thank you very much.