

به نام خدا

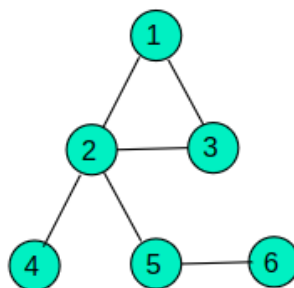
امتحان میان ترم درس گراف کاوی

نکات مهم

- صورت سوالات را با دقت بخوانید و پاسخ ها را بدون ابهام و دقیق در برگه پاسخ بنویسید
 - از هر برگه برای پاسخ دهی به تنها یک سوال استفاده کنید
 - در صورت نیاز می توانید از نمودارها یا مثال ها برای پشتیبانی از پاسخ های خود استفاده کنید
- با آرزوی موفقیت

سوال ۱

گراف زیر را به عنوان یک شبکه اجتماعی از کاربران و تعاملات آنها در نظر بگیرید



الف) معیارهای زیر را برای هر کاربر در این شبکه اجتماعی محاسبه کنید

degree

closeness centrality

betweenness centrality

کدام کاربران دارای بیشترین و کمترین مقادیر هر یک از این معیارها هستند؟ هر یک از مقادیر فوق در به کارگیری این معیارها در مهندسی ویژگی برای الگوریتم های یادگیری ماشین که رفتار شبکه اجتماعی یا تعامل کاربر را پیش بینی می کنند چگونه تفسیر می شوند؟ (کاربران را از این نظر با یکدیگر مقایسه کنید)

ب) معیارهای زیر را برای هر کاربر در شبکه اجتماعی محاسبه کنید

clustering coefficient

Graphlets of size 3 vector

هر یک از مقادیر فوق چگونه تفسیر می‌شوند؟ کاربران را از این نظر با یکدیگر مقایسه کنید

(ج) فرض کنید می‌خواهید با هدف تبلیغات هدفمند، تاثیرگذارترین کاربر در شبکه اجتماعی را شناسایی کنید. کدام کاربر بیشترین تأثیر را بر معیارهای تعامل شبکه خواهد داشت و چرا؟ (بر اساس معیارهای مطرح شده در درس پاسخ دهید)

(د) تعاملات (یال‌های) موجود در شبکه اجتماعی مورد اشاره را از نظر اهمیت با استفاده از حداقل ۳ شاخص با یکدیگر مقایسه کنید. کدام تعامل (تعاملات) از بقیه با اهمیت‌تر هستند؟ (در مورد تفسیر شاخص‌ها به‌طور دقیق توضیح دهید)

سوال ۲

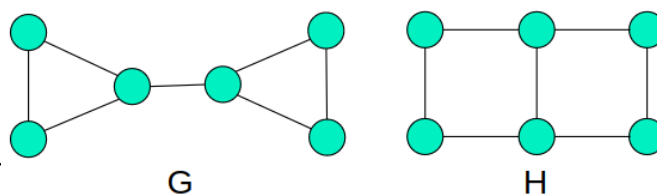
(الف) ایده اصلی الگوریتم نود۲وک^۱ و تفاوت آن با دیپ‌واک^۲ را توضیح دهید. مقادیر مختلف برای پارامترهای^۳ الگوریتم نود۲وک چگونه تفسیر می‌شوند؟ (برای پاسخ به بخش دوم سوال می‌توانید از مثال‌های خاص برای مقادیر پارامترها استفاده کنید)

(ب) نحوه انتخاب قدم زدن تصادفی^۴ برای گراف شبکه اجتماعی سوال ۱ را با استفاده از الگوریتم نود۲وک با یک مثال توضیح دهید (راهنمایی: یک کاربر را به دلخواه به عنوان راس ابتدایی در نظر بگیرید و با در نظر گرفتن مقادیر دلخواه انتخابی خود برای پارامترهای الگوریتم، نحوه انتخاب یک قدم زدن تصادفی را به‌طور دقیق توضیح دهید)

(ج) دو الگوریتم نود۲وک و دیپ‌واک را از نظر نقاط قوت و ضعف با هم مقایسه کنید. کدام الگوریتم را برای یک مجموعه داده گرافی پیشنهاد می‌کنید و چرا؟

سوال ۳

مسئله یکریختی گراف‌ها یک مسئله بسیار دشوار در ریاضیات است. الگوریتم وایسفالر-لیمن به‌عنوان یکی از قدرتمندترین روش‌ها برای تشخیص عدم یکریختی دو گراف شناخته می‌شود. دو گراف غیریکریخت زیر را در نظر بگیرید و به سوالات پاسخ دهید



- 1 node2vec
- 2 DeepWalk
- 3 p and q parameters
- 4 random walk

الف) آیا الگوریتم وایسفایلر-لیمن قادر به تشخیص عدم یکریختی دو گراف بالا هست؟ (چرایی پاسخ خود را به طور دقیق با شکل توضیح دهید)

ب) فرض کنید که همه راس‌ها بردارهای ویژگی اولیه یکسانی دارند و ما از یک نسخه ساده شده از شبکه عصبی گرافی استفاده می‌کنیم (بدون تابع فعال‌سازی و بدون شبکه عصبی). همچنین رابطه زیر برای محاسبه فیچر هر راس در هر لایه از شبکه عصبی گرافی در نظر گرفته می‌شود

$$h_v^l = \text{Mean} \left(\left\{ h_u^{l-1} \right\}_{u \in N(v)} \right) + h_v^{l-1}$$

آیا شبکه عصبی گرافی با تعریف بالا قادر به تشخیص عدم یکریختی دو گراف بالا هست؟ (چرایی پاسخ خود را به طور دقیق با شکل توضیح دهید)

ج) با به‌کارگیری یکی از روش‌های مطرح شده در درس (روشنی غیر از دو روش مطرح شده در قسمت‌های الف و ب و همچنین بدون استفاده از روش‌های معمول ریاضیاتی)، نشان دهید که دو گراف بالا یکریخت نیستند

سوال دارای امتیاز اضافه: شبکه‌های عصبی گرافی پیچشی^۵ و دو روش بندهای الف و ج را از نظر نقاط قوت و ضعف با یکدیگر مقایسه کنید (حداقل ۳ مورد مقایسه بیان شود)