



پیش‌بینی پیوند

پروژه درس شبکه‌های پیچیده و پویا

تاریخ	۱۴۰۲/اسفند/۲۹
شماره پروژه	۱
مهلت انجام	۱۴۰۳/فروردین/۲۱

❖ قبل از انجام دادن پروژه به موارد زیر توجه کنید.

نکته	
۱	تحويل پروژه با تأخير موجب كسر نمره مي‌گردد.
۲	جواب‌ها از لحاظ شباهت بررسی می‌شوند و جواب‌های مشابه مورد قبول نیستند.
۳	تمامی کدها باید به همراه فایل گزارش ارسال شوند.
۴	کسانی که از روش‌های یادگیری ماشین استفاده کنند نمره اضافی دریافت خواهند کرد.
۵	با انتخاب یک مقاله که در چارچوب این پروژه قرار می‌گیرد و پیاده‌سازی الگوریتم پیشنهادی نیز می‌توان نمره اضافی دریافت کرد.
۶	خلاقیت نمره اضافی دارد.

❖ شناسایی پیوندهای گمشده و پیش‌بینی پیوندهای آینده یکی از شاخه‌های مهم تحلیل شبکه‌های اجتماعی است. پیش‌بینی پیوند به عنوان تخمین احتمال تشکیل پیوند بین هر جفت گره که پیوندی برای آنها وجود ندارد تعریف می‌شود. داده‌های استفاده شده در این پروژه اطلاعات شبکه‌ی فیسبوک است که تنها شامل اطلاعات پیوندها می‌شود. داده‌ها به عنوان فایل پیوست ارسال شده است. هدف پیش‌بینی پیوند و پیشنهاد یک گره به هر گره در شبکه با عنوان "فردی که ممکن است بشناسید" است. موارد زیر را انجام دهید و خروجی هر بخش را گزارش کنید.

ردیف	
	خود نظارتی
۱	پیش‌بینی پیوند معمولاً کاری بدون نظارت ^۱ یا با نظارت خود ^۲ است، به این معنی که گاهی اوقات نیاز است که خودمان برچسب‌های مربوطه را ایجاد کنیم. برای خود نظارتی، از بین یال‌های متصل به هر گره در صورت امکان یک پیوند را به صورت تصادفی حذف کرده و این پیوندها را گزارش کنید. پیوندهای حذف شده در مرحله‌ی ۳ برای آزمودن دقت الگوریتم شما استفاده می‌شود. لازم به ذکر است که کتابخانه‌هایی برای جداسازی داده‌ها به داده‌های آموزش و تست وجود دارد که می‌توانید از آنها نیز استفاده کنید.
	پیش‌بینی پیوند
۲	رویکرد اصلی پیش‌بینی پیوند در ساده‌ترین حالت به شرح زیر است:

¹ Unsupervised

² Self-supervised

<p>(۱) با پیوندهای باقی مانده گراف بسازید و الگوریتمی را برای محاسبه‌ی ماتریس شباهت برای هر جفت گره در گراف پیشنهاد کنید.</p> <p>(۲) برای محاسبه‌ی شباهت بین دو گره در شبکه، می‌توانید از انواع معیارهای شباهت مثل ضریب ژاکار^۳ (تعداد همسایگان مشترک را برای محاسبه شباهت بین جفت گره‌ها در نظر می‌گیرد)، آدامیک آدار^۴ (بر اساس پیوندهای مشترک بین دو گره شباهت دو گره را محاسبه می‌کند) و غیره استفاده کنید. اگر از روش شباهت استفاده می‌کنید حداقل ۴ معیار شباهت انتخاب کرده و دلیل خود را برای انتخاب این معیارها گزارش دهید.</p> <p>(۳) از ماتریس شباهت نهایی را گزارش کنید.</p> <p>(۴) از ماتریس شباهت به دست آمده استفاده کنید تا برای هر گره گره‌ای را به عنوان "فردی که ممکن است بشناسید" توصیه کنید.</p>	
آزمون	
پیوندهای پیش‌بینی شده را با پیوندهای حذف شده مقایسه کرده و خروجی را گزارش کنید.	۳

³ Jaccard Coefficient

⁴ Adamic Adar