



**آنچه باید در تحویل داده شود:** شما می‌بایست برای هر پروژه دو فایل تحویل دهید، یکی فایل کامل شده jupyter-notebook است و دیگری گزارشی مناسب در قالب PDF است. توجه داشته که فایل نوت‌بوک باید فایلی باشد که گزارش‌های آموزش شبکه و شکل‌ها و تصاویر به‌طور کامل در آن باشد و در صورت نیاز توضیحاتی درباره چگونگی پیاده‌سازی در بالای هر سلول نوشته‌شده باشد.

این پروژه‌ها به‌صورت انفرادی انجام و تحویل داده می‌شوند. در گزارش PDF به‌طور جامع‌تری درباره مسئله و چگونگی راه‌حل پیشنهادشده توضیح دهید و نیز سؤالاتی را که در پایین هر مسئله آورده شده است پاسخ دهید. در ضمن چنانچه از مقاله و یا سایت‌های مختلف برای راه‌حل استفاده می‌کنید بدیهی است که نیاز به ارجاع دارد، سعی کنید منابعی را که در حل سؤال استفاده می‌کنید را در انتهای گزارش معرفی کنید.

## پروژه ۲: تولید اشعار شاهنامه با استفاده از یک شبکه بازگشتی.

در این پروژه با استفاده از داده‌های متن شاهنامه که در قالب یک فایل text در اختیاران گذاشته شده است، می‌خواهیم که یک شبکه را آموزش دهید به طوری که وقتی یک دنباله از بیت‌های شاهنامه را بدهیم بتوانید، آن را ادامه دهد. توجه شود که قرار نیست دقیقاً همان بیت‌ها را تولید کنید و یا به طور دقیق‌تر هدف تولید یک جدول جستجو<sup>۱</sup> نیست و جملات می‌بایست جدید باشند.

## پیش‌نیازهای پروژه ۲:

برای این پروژه هم پیشنهاد می‌شود که با پایتون ۳ کار کنید چرا که نسبت به حروف فارسی و یونیکد سازگاری بیشتری دارد. همچنین به سایر کتابخانه‌هایی که در پروژه ۱ گفته شد بجز sklearn و matplotlib نیاز خواهید داشت، هر چند که ممکن است در راه حل پیشنهادیتان، از این کتابخانه‌ها و یا بسته‌های دیگری نیز استفاده کنید.

همچنین اگر در راه حل خود فکر می‌کنید که نیاز به اطلاعات و ابزارهای دیگری دارید که دسترسی به آن‌ها راحت نیست، بهتر است ابتدا به دنبال راه حل شدنی‌تر بگردید و در غیر این صورت بهتر است پیش از ادامه دادن با حل تمرین یا استادتان مشورت کنید. بدیهی است که در این صورت انتظار ایده‌های بهتری را برای شبیه‌سازی خواهیم داشت.

<sup>1</sup>Lookup table.

## پایگاه داده پروژه ۲:

همانطور که عنوان شد این پایگاه داده به صورت یک فایل خام متنی (با پسوند txt) آماده شده و قابل استفاده است.

## سؤالات مربوط به پروژه ۲:

۱- این مسأله را می‌توان در دو سطح واژه و حرف پیاده‌سازی کرد، اول آنکه توضیح دهید هر مورد را که ساختار داده‌ها و معماری شبکه به چه صورت تغییر می‌کند. دوم آنکه به نظر شما کدام مدل به نتیجه بهتر می‌رسد و کدام مدل ساده‌تر پیاده‌سازی می‌شود؟ چرا؟

۲- شبکه‌های بازگشتی می‌توانند در دو حالت یک جهتی و یا دو جهتی<sup>۲</sup> طراحی شوند، به نظر شما در این پروژه کدامیک بهتر جواب می‌دهد؟ (اگر امکانش هست آن را به صورت عملی نشان دهید).

۳- فرابرازش در حل این مسأله به چه صورت نمایان می‌شود؟ راه حل شما برای مقابله با آن چیست؟

۴- در گزارش خود به طور دقیق توضیح دهید که داده‌ها را چگونه برای شبکه آماده کردید.

۵- آیا مدل شما می‌تواند یک شاهنامه با معنی و جدید بسازد؟ چه ضعف‌هایی دارد و چرا؟ بد نیست با نگاهی مختصر به مقالات، در مورد ایده‌های آن‌ها به طور مختصر توضیحاتی ارایه دهید.

در اینترنت منابع مختلفی را می‌توانید برای راهنمایی پیدا کنید، توجه کنید که این مسأله هم به روش‌های مختلفی قابل پیاده‌سازی است، بنابراین سعی کنید از راهنمایی‌های مختلف استفاده کنید ولی در نهایت ایده‌های خودتان را ادامه دهید و در گزارش توضیح دهید.

برای نمونه در سایت زیر راهنمایی ساده و مقدماتی ارایه شده است:

<https://www.kaggle.com/shivamb/beginners-guide-to-text-generation-using-lstms>

هر چند که تنها روند پیاده‌سازی برای این مسأله نیست.

با تشکر.

موفق باشید.

---

<sup>2</sup>Bidirectional.