



بازیابی پیشرفته اطلاعات

نیم‌سال اول ۱۴۰۲-۰۳

مدرس: دکتر حمید بیگی

زمان: ۲۴ مهر

کوئیز

سری اول (۱۰۰ نمره)

سوال ۱: پاسخ کوتاه (۳۵ نمره)

برای سوالات زیر پاسخ کوتاه بنویسید.

- نقش idf در ساخته شدن بردار tf-idf چیست؟

- تفاوت Stemming و Lemmatizaion در چیست؟

- ، مطابق الگوی مبتنی بر برنامه‌ریزی پویا اگر هزینه‌ی افزایش و کاهش برابر ۱ و هزینه‌ی تغییر برابر ۳ باشد کمترین هزینه کلمات زیر را بدست آوردید.

- فرسایش - آسایش

- تقطیر - تغییر

سوال ۲: Sorting Algorithm (۲۰ نمره)

External Sorting Algorithm را توضیح داده و ۲ مشکل آن را بیان کنید.

سوال ۳: اسناد (۴۵ نمره)

۳ سند زیر را در نظر بگیرید که پس از عملیات پیش پردازش به صورت زیر خواهند بود:

داک ۱: learning mechanics build machine

داک ۲: artificial intelligence machine learning

داک ۳: text mining deep learning

- شاخص معکوس (inverted index) را برای کلمه‌های learning mechanics build machine، deep، رسم کنید.
- برای کلمه‌های learning mechanics build machine، deep، ماتریس term-document را رسم کنید.
- با توجه به ماتریس رسم شده اسنادی که با پرسش "machine learning" مرتبط هستند را بدست آورید.