

گزارش فاز اول پژوهشی بازیابی پیشرفته‌ی اطلاعات

سحر موالی

غزل ابراهیمی

بخش اول

در این بخش، الگوریتم‌های دسته‌بندی Naive Bayes، KNN، SVM و Random Forest را پیاده‌سازی کردیم و با اجرای آن‌ها روی داده‌های آزمون، نتیجه را با دسته‌بندی داده‌شده مقایسه کردیم.

بخش دوم

در این بخش، الگوریتم جستجو با استفاده از tf-idf را که در فاز قبل پیاده‌سازی کرده بودیم، با استفاده از الگوریتم‌های نوشته شده در بخش قبل تغییر دادیم. به این صورت کهتابع جستجو علاوه بر عبارت مورد جستجو شماره‌ی دسته را هم به عنوان وردی دریافت کرده و عبارت را تنها در اسناد آن دسته جستجو می‌کند. اسناد را هم با استفاده از Naïve Bayes دسته‌بندی می‌کنیم.

بخش سوم

مقادیر خواسته شده در جدول زیر آمده‌اند: (دقیق کم نتایج به دلیل اجرای الگوریتم بر روی تعداد کمی از داده‌های آموزش است.)

Algorithm	Accuracy	Precision	Recall	F1-score
Naive Bayes	0.42	0.56	0.83	0.67
KNN	0.63	0.5	0.67	0.55
SVM	0.32	0.32	1	0.48
Random Forest	0.32	0.44	0.67	0.53

مشارکت اعضای گروه در انجام پژوهه

در انجام این فاز پژوهه هر یک از اعضای گروه به میزان ۵۰ درصد مشارکت داشتند.