 **بازیابی پیشرفته اطلاعات**

مدرس: دکتر بیگی

شماره گروه: 5

دانشکده مهندسی کامپیوتر تهیه‌کنندگان: نیما جمالی – سپهر فعلی – سینا کاظمی

گزارش فاز اول پروژه

فهرست مطالب

[1- پیش‌پردازش اولیه 3](#_Toc54662803)

[پیش‌پردازش مستندات انگلیسی 3](#_Toc54662804)

[کلمات پرتکرار مستندات انگلیسی 3](#_Toc54662805)

[پیش‌پردازش مستندات فارسی 3](#_Toc54662806)

[کلمات پرتکرار مستندات فارسی 3](#_Toc54662807)

[2- نمایه‌سازی 3](#_Toc54662808)

[نمایه bigram 3](#_Toc54662809)

[نمایه positional 3](#_Toc54662810)

[توابع درج و حذف مستندات 4](#_Toc54662811)

[توابع ذخیره‌سازی و لود نمایه‌ها 4](#_Toc54662812)

[3- فشرده‌سازی نمایه‌ها 4](#_Toc54662813)

[فشرده‌سازی با variable byte 4](#_Toc54662814)

[فشرده‌سازی با gamma code 4](#_Toc54662815)

[ذخیره‌سازی و لود نمایه‌ها 4](#_Toc54662816)

[4- اصلاح پرسمان 4](#_Toc54662817)

[استفاده از نمایه bigram و معیار جاکارد برای یافتن کلمات مشابه 5](#_Toc54662818)

[تابع محاسبه‌ی فاصله ویرایش دو کلمه 5](#_Toc54662819)

[نمایش پرسمان اصلاح‌شده 5](#_Toc54662820)

[5- جستجو و بازیابی اسناد 5](#_Toc54662821)

[سر تیتر اول 5](#_Toc54662822)

[سر تیتر دوم 5](#_Toc54662823)

[6- نحوه‌ی تقسیم وظایف 5](#_Toc54662824)

[نیما جمالی 5](#_Toc54662825)

[سپهر فعلی 6](#_Toc54662826)

[سینا کاظمی 6](#_Toc54662827)

[7- مراجع 6](#_Toc54662828)

# پیش‌پردازش اولیه

توضیحات با استایل doc style

## پیش‌پردازش مستندات انگلیسی

متن

## کلمات پرتکرار مستندات انگلیسی

متن

## پیش‌پردازش مستندات فارسی

متن

## کلمات پرتکرار مستندات فارسی

متن

# نمایه‌سازی

توضیحات

## نمایه bigram

متن

## نمایه positional

متن

## توابع درج و حذف مستندات

متن

## توابع ذخیره‌سازی و لود نمایه‌ها

متن

# فشرده‌سازی نمایه‌ها

توضیحات

## فشرده‌سازی با variable byte

متن

## فشرده‌سازی با gamma code

متن

## ذخیره‌سازی و لود نمایه‌ها

متن

# اصلاح پرسمان

توضیحات

## استفاده از نمایه bigram و معیار جاکارد برای یافتن کلمات مشابه

متن

## تابع محاسبه‌ی فاصله ویرایش دو کلمه

متن

## نمایش پرسمان اصلاح‌شده

متن

# جستجو و بازیابی اسناد

توضیحات

## سر تیتر اول

متن

## سر تیتر دوم

متن

# نحوه‌ی تقسیم وظایف

توضیحات

## نیما جمالی

1-

2-

3-

## سپهر فعلی

1-

2-

3-

## سینا کاظمی

1-

2-

3-

# مراجع