

# پروژه پایانی درس «مبانی بازیابی اطلاعات و جستجوی وب»

نیمسال ۱-۹۹-۹۸

## هدف: طراحی یک سیستم کوچک جستجوی اطلاعات راجع به دانشگاه آزاد اسلامی واحد کرج از سایت ویکی‌پدیا

صفحه دانشگاه آزاد اسلامی واحد کرج در سایت ویکی‌پدیا:

[https://fa.wikipedia.org/wiki/%D8%AF%D8%A7%D9%86%D8%B4%DA%AF%D8%A7%D9%87\\_%D8%A2%D8%B2%D8%A7%D8%AF\\_%D8%A7%D8%B3%D9%84%D8%A7%D9%85%D8%8C\\_%D9%88%D8%A7%D8%AD%D8%AF\\_%DA%A9%D8%B1%D8%AC](https://fa.wikipedia.org/wiki/%D8%AF%D8%A7%D9%86%D8%B4%DA%AF%D8%A7%D9%87_%D8%A2%D8%B2%D8%A7%D8%AF_%D8%A7%D8%B3%D9%84%D8%A7%D9%85%D8%8C_%D9%88%D8%A7%D8%AD%D8%AF_%DA%A9%D8%B1%D8%AC)

W ویکی

+

دانشگاه آزاد اسلامی واحد کرج - ویکی

← → ↺

fa.wikipedia.org/wiki/دانشگاه\_آزاد\_اسلامی\_واحد\_کرج

🔍 ⚙️ 🌐 📄

👤

👤 به سامانه وارد نشده‌اید بحث مشارکت‌ها ایجاد حساب کاربری ورود

🔍 جستجو در ویکی‌پدیا

📄 مقاله

📄 بحث

📄 ویرایش

📄 نمایش تاریخچه

## دانشگاه آزاد اسلامی واحد کرج

از ویکی‌پدیا، دانشنامهٔ آزاد

مختصات: ۳۵°۸۵′۳۵″شمالی ۵۰°۹′۵۰″شرقی﻿ / ﻿۳۵٫۹۲۶۳۹°شمالی ۵۰٫۱۶۳۸۹°شرقی﻿ / ۳۵.۹۲۶۳۹; ۵۰.۱۶۳۸۹

**دانشگاه آزاد اسلامی واحد کرج**

دانشگاه آزاد اسلامی

آرامان ایرانی برای جهانی شستن

**مستحضات واحد**

بنیانگذاری

۱۳۶۳

شهر

کرج

رئیس

محمد حسن برهانی فر

معاون آموزشی

حسن شونیدی

معاون پژوهشی

سید احسان شیرنگی

معاون دانشجویی

ابراهیم اکبری

هیئت علمی

۷۰۰ (در کل استان البرز)

کارمندان

۹۰۰ (در کل استان البرز)

دانشجویان

۲۵۰۰۰ (در کل استان البرز)

کاردانی

۵۰۰۰

کارشناسی

۳۰۰۰۰

تحصیلات تکمیلی

۱۰۰۰۰

کارشناسی ارشد

۹۰۰۰

دکترا

۱۰۰۰

دانش‌آموختگان

۸۰۰۰۰ (در کل استان البرز)

فضای آموزشی

۴۷ نانو متر مکعب

نوع واحد

جامع الف

منطقه

شهری

وبگاه

http://www.kiau.ac.ir

نمایی از یکی از خیابان‌ها و مسجد دانشگاه آزاد اسلامی واحد کرج

**محتویات [تفہن]**

۱ پیشینه

۲ دانشکده‌ها، مراکز و گروه‌های آموزشی مستقل واحد کرج

۳ رشته‌ها، گرایش‌ها و مقاطع تحصیلی

۳.۱ دکتری تخصصی

۳.۲ دکتری حرفه‌ای

۳.۳ کارشناسی ارشد

۳.۴ کارشناسی پیوسته

۳.۵ کارشناسی ناپیوسته

۳.۶ کاردانی ناپیوسته

۳.۷ کاردانی پیوسته

۴ امکانات آموزشی

۴.۱ کارگاه‌ها و آزمایشگاه‌ها

۴.۲ کتابخانه‌ها

۴.۳ سایت رایانه‌ای

۴.۴ لابراتوارهای زبان انگلیسی و عربی

۵ فعالیت‌های فوق‌برنامه

۶ فعالیت‌های پژوهشی

۶.۱ انجمن‌های علمی، فرهنگی، هنری

۶.۲ نشریات دانشجویی واحد کرج

۶.۳ باشگاه پژوهشگران جوان

۶.۴ سمینارهای علمی

۷ تسهیلات رفاهی

## شرح پروژه:

با استفاده از «ابزار هضم»<sup>۱</sup>، مجموعه دستورات و برنامه‌های لازم، سیستم بازیابی اطلاعاتی طراحی کنید که توکن‌ها (نشانه‌های) موجود در یک صفحه فوق که صفحه «دانشگاه آزاد اسلامی واحد کرج» در ویکی‌پدیا است را استخراج کرده، پس از انجام عملیات ریشه‌گیری کلمات فارسی، یک ساختار شاخص موقعیتی (Positional Index) بسازد.

این سیستم باید امکان جستجو با عملگرهای AND و OR و جستجوی شامل عبارات دقیقی مانند "مهندسی کامپیوتر" و "نشریات دانشگاهی" در این شاخص موقعیتی را داشته باشد. برای سادگی، در هر عبارت پرس‌وجو، حداکثر یک عملگر AND یا OR استفاده می‌شود و اگر عملگر "" (Double Quotation) استفاده شود، عملگرهای AND و OR در پرس‌وجو مجاز نیستند.

نشان دهید که آیا عبارات زیر در این سند وجود دارند یا خیر:

عبارت پرس‌وجو	نتیجه
"مهندسی کامپیوتر"	بله (به همراه توکن‌های چپ و راست)
"فیش هفتگی"	بله (به همراه توکن‌های چپ و راست)
هیچ AND دانشگاه	خیر
حضور AND اقتصاد	بله
هیچ OR دانشگاه	بله
"مجاز به انتشار آن به"	بله (به همراه توکن‌های چپ و راست)
سمینار	بله (توضیح: در متن بصورت جمع آمده و طی ریشه‌گیری به مفرد تبدیل شده است)

در صورتی که عبارت پرس‌وجو شامل عملگر "" (Double Quotation) باشد، در کنار نتیجه‌ی نمایش داده شده، یک توکن از سمت چپ و یک توکن از سمت راست از داخل ساختار شاخص (و نه از درون سند) باید استخراج شده و نمایش داده شود. برای مثال، در صورت جستجوی عبارت "فیش هفتگی"، علاوه بر اینکه نتیجه «بله» باید نمایش داده شود، توکن‌های «تهیه» و «از» نیز باید به شیوه مناسبی نمایش داده شود، زیرا عبارت «تهیه فیش هفتگی از» در متن این سند هست.

پرینت متن برنامه و هفت تصویر (Screenshot) از نتیجه اجرای برنامه خودتان برای هفت پرس‌وجوی جدول فوق را حداکثر تا پایان آذرماه ۱۳۹۸ تحویل دهید و همچنین متن برنامه و تصاویر را به آدرس [hsaboohi@gmail.com](mailto:hsaboohi@gmail.com) ایمیل کنید. (توجه بفرمایید: پرینت و ایمیل، هر دو لازم هستند)

موفق باشید

هادی صبحی

<sup>1</sup> <http://www.sobhe.ir/hazm/>