بسم الله الرحمن الرحیم

هیرو کریم پور

پروژه معماری

Contents

[تعریف مسئله: 3](#_Toc13240934)

[ابر کلمات : 3](#_Toc13240935)

[میزان رضایت کاربران توییتر: 3](#_Toc13240936)

[محبوب ترین کاربران اینستا گرام 4](#_Toc13240937)

[بالاترین منابع منتشر کننده خبر 4](#_Toc13240938)

[پربازدیدترین کانال ها در تلگرام 4](#_Toc13240939)

[ابرموضوعات خبر : 5](#_Toc13240940)

[تعداد اخبار منتشر شده : 5](#_Toc13240941)

[سر خط خبر ها : 5](#_Toc13240942)

[صفحه گزارشات اینستاگرام 6](#_Toc13240943)

[صفحه گزارشات توییتر 7](#_Toc13240944)

[صفحه گزارشات تلگرام 9](#_Toc13240945)

[فرم جستجو در سامانه 11](#_Toc13240946)

[سناریوهای غیروظیفه مندی: 11](#_Toc13240947)

[ ایجاد glossary : تعریف برخی اصطلاحات 12](#_Toc13240948)

[ قالب کلی درخواست‌های وب سرویس 14](#_Toc13240949)

[ مشاهده تمام گزارش‏ها موجود در وب‏سرویس 14](#_Toc13240950)

# 

# تعریف مسئله:

در این پروژه داشبوردی طراحی شده که در آن سایت های خبری و اینستاگرام و تویتر و ... کرال شده و اطلاعات در داشبورد تعبیه می شود . کاربر در ابتدا وارد صفحه ورود به پنل کاربری می شود برای هر کاربر نام کاربری و رمز عبور تعیین شده است کاربر نام کاربری و رمز عبور خود را وارد می کندسیستم ابتدا کاربر را اعتبار سنجی می کند با صحیح بودن نام کاربری و رمز عبور به داشبورد سیستم انتقال داده می شود.

داشبورداصلی سیستم شامل منوی می باشد که شامل ایتم های از قبیل خانه ، جستجو ، اخبار ، توییتر ، اینستاگرام ، تلگرام ، ارتباط با ما ، راهنما و خروج از سیستم می باشد.

صفحه اصلی داشبورد شامل گزارش های متعددی از قبیل ابر هشتگ های تلگرام ، ابر کلمات اخبار ، میزان رضایت کاربران توییتر ، آخرین اخبار منتشر شده ، محبوب ترین توییت های منتشر شده ، محبوبترین کاربران اینستا گرام ، بالاترین منتشر کنندگان خبر ، روند فراوانی انتشار توییتر ، پربازدید ترین کانال های تلگرام نشان داده شده است.

# ابر کلمات :

در تمام گزارشات ، ابر کلمات یا ابر هشتگ ها اطلاعاتی با ویژگی های از قبیل عنوان ابر ، عنوان کلمه مورد نظر و وزن کلمه نشان داده شده است در زیر هر ابرکلمات یک فیلتر براساس تاریخ نیز قرار داده شده که شما می توانید خروجی ابر کلمات را از یک تاریخ تا یک تاریخ مشخص فیلتر نمایید و یک دکمه لیست نیز در زیر ابر کلمات قرار داده شده که می توانید خروجی را به صورتی لیستی مشاهده نماید.

نیازمندیهای وظیفه مندی:

# میزان رضایت کاربران توییتر:

در این گزارش عنوان گزارش میزان مخالف ، خنثی ، موافق مطالب براساس درصد و تعداد نشان داده می شود.

گزارش خبر ها:

در این نوع گزارش عنوان گزارش ، عنوان خبر ، خلاصه خبر ، لینک خبر ، تصویر خبر ، تاریخ خبر ، منبع خبر نشان داده شده می شود در زیر گزارش یک فیلتر براساس تاریخ نیز قرار داده شده که شما می توانید خروجی را از یک تاریخ تا یک تاریخ مشخص فیلتر نمایید.

گزارش از توییت ها:

در این نوع گزارش عنوان توییت ، تعداد لایک ، تعداد تریتویت ، تاریخ ، نام کاربری نشان داده می شود در زیر گزارش یک فیلتر براساس تاریخ نیز قرار داده شده که شما می توانید خروجی را از یک تاریخ تا یک تاریخ مشخص فیلتر نمایید.

# محبوب ترین کاربران اینستا گرام

در این نوع گزارش تصویر کاربر ، عنوان ، نام کاربری و تعداد لایک نشان داده می شود در زیر گزارش یک فیلتر براساس تاریخ نیز قرار داده شده که شما می توانید خروجی را از یک تاریخ تا یک تاریخ مشخص فیلتر نمایید.

# بالاترین منابع منتشر کننده خبر

در این نوع گزارش تصویر منبع ، عنوان ، نام کاربری و تعداد خبر نشان داده می شود در زیر گزارش یک فیلتر براساس تاریخ نیز قرار داده شده که شما می توانید خروجی را از یک تاریخ تا یک تاریخ مشخص فیلتر نمایید.

# پربازدیدترین کانال ها در تلگرام

در این نوع گزارش عنوان ، نام کاربری و تعداد مشاهده به صورت نمودارنشان داده می شود در زیر گزارش یک فیلتر براساس تاریخ نیز قرار داده شده که شما می توانید خروجی را از یک تاریخ تا یک تاریخ مشخص فیلتر نمایید.

صفحه اخبار

صفحه اخبار شامل گزارش های زیر می باشد.

# ابرموضوعات خبر :

درابرموضوعات اخبار اطلاعاتی با ویژگی های از قبیل عنوان ابر ، عنوان کلمه مورد نظر و وزن کلمه نشان داده شده است در زیر هر ابرکلمات یک فیلتر براساس تاریخ نیز قرار داده شده که شما می توانید خروجی ابر کلمات را از یک تاریخ تا یک تاریخ مشخص فیلتر نمایید و یک دکمه لیست نیز در زیر ابر کلمات قرار داده شده که می توانید خروجی را به صورتی لیستی مشاهده نماید.

# تعداد اخبار منتشر شده :

در این گزارش تعداد خبر و نرخ توزیع خبر نشان داده می شود.

تعداد اخبار منتشر شونده :

در این نوع گزارش تعداد اخبار منتشر شده در هفته اخیر،30 روز گذشته ، یک ساعت قبل و 24 ساعت گذشته نشان داده می شود.

# سر خط خبر ها :

در این نوع انواع گزارش از قبیل آخرین اخبار ، پربحث ترین خبر ، پربازنشرترین خبر ، پرنفوذترین اخبار نشان داده شده است. در این نوع گزارش ها عنوان گزارش ، عنوان خبر ، خلاصه خبر ، لینک خبر ، تصویر خبر ، تاریخ خبر ، منبع خبر نشان داده شده می شود در زیر گزارش یک فیلتر براساس تاریخ نیز قرار داده شده که شما می توانید خروجی را از یک تاریخ تا یک تاریخ مشخص فیلتر نمایید.

منابع دارای بیشترین تعداد اخبار :

در این نوع گزارش تصویر منبع ، عنوان ، نام کاربری و تعداد خبر نشان داده می شود در زیر گزارش یک فیلتر براساس تاریخ نیز قرار داده شده که شما می توانید خروجی را از یک تاریخ تا یک تاریخ مشخص فیلتر نمایید.

نمودار تعداد خبر

یک نمودار کامل از تعداد اخبار براساس تاریخ به شما نشان می دهد در زیر گزارش یک فیلتر براساس تاریخ نیز قرار داده شده که شما می توانید خروجی را از یک تاریخ تا یک تاریخ مشخص فیلتر نمایید.

# صفحه گزارشات اینستاگرام

ابرهشتگ های اینستاگرام ، ابرکلمات اینستاگرام :

در ابرهشتگ های اینستاگرام ، ابرکلمات اینستاگرام اطلاعاتی با ویژگی های از قبیل عنوان ابر ، عنوان کلمه مورد نظر و وزن کلمه نشان داده شده است در زیر هر ابرکلمات یک فیلتر براساس تاریخ نیز قرار داده شده که شما می توانید خروجی ابر کلمات را از یک تاریخ تا یک تاریخ مشخص فیلتر نمایید و یک دکمه لیست نیز در زیر ابر کلمات قرار داده شده که می توانید خروجی را به صورتی لیستی مشاهده نماید.

گزارشات کلی اینستاگرام پست ها

در این بخش پست های محبوب ، پست های بربحث ، پست های پرنفوذ ، آخرین پست نمایش داده می شود.در این گزارشات اطلاعاتی از قبیل نام کاربر ، تصویرکاربر، عنوان کاربر ، تعداد لایک ، تعداد کامنت ، و توضیح گزارش نشان داده می شود در زیر گزارش یک فیلتر براساس تاریخ نیز قرار داده شده که شما می توانید خروجی را از یک تاریخ تا یک تاریخ مشخص فیلتر نمایید.

گزارشات کلی اینستاگرام کاربران

در این بخش کاربرهای پربحث ، بیشترین منتشرکنندگان ، محبوبترین کاربران ، پرنفوذترین کاربران نمایش داده می شود.در این گزارشات اطلاعاتی از قبیل نام کاربر ، تصویرکاربر، عنوان کاربر ، تعداد لایک ، تعداد کامنت ، و توضیح گزارش نشان داده می شود در زیر گزارش یک فیلتر براساس تاریخ نیز قرار داده شده که شما می توانید خروجی را از یک تاریخ تا یک تاریخ مشخص فیلتر نمایید.

نمودار انتشار پست

یک نمودار کامل از نمودار انتشار پست براساس تاریخ به شما نشان می دهد در زیر گزارش یک فیلتر براساس تاریخ نیز قرار داده شده که شما می توانید خروجی را از یک تاریخ تا یک تاریخ مشخص فیلتر نمایید.

# صفحه گزارشات توییتر

ابرهشتگ های توییتر ، ابرکلمات توییتر:

در ابرهشتگ های توییتر ، ابرکلمات توییتر اطلاعاتی با ویژگی های از قبیل عنوان ابر ، عنوان کلمه مورد نظر و وزن کلمه نشان داده شده است در زیر هر ابرکلمات یک فیلتر براساس تاریخ نیز قرار داده شده که شما می توانید خروجی ابر کلمات را از یک تاریخ تا یک تاریخ مشخص فیلتر نمایید و یک دکمه لیست نیز در زیر ابر کلمات قرار داده شده که می توانید خروجی را به صورتی لیستی مشاهده نماید.

تعداد کاربران ونرخ توزیع توییت :

در این گزارش تعداد کاربران و نرخ توزیع خبر به صورت کامل نشان داده می شود.

میزان رضایت کاربران توییتر:

در این گزارش عنوان گزارش میزان مخالف ، خنثی ، موافق مطالب براساس درصد و تعداد نشان داده می شود.

سر خط توییت ها :

در این نوع انواع گزارش از قبیل محبوب ترین ، تاثیرگزارترین ، آخرین ، پرنازنشترین نشان داده شده است. در این نوع گزارش ها عنوان گزارش ، عنوان تویت ، خلاصه تویت ، لینک تویت ، تاریخ تویت ، منبع تویت نشان داده شده می شود در زیر گزارش یک فیلتر براساس تاریخ نیز قرار داده شده که شما می توانید خروجی را از یک تاریخ تا یک تاریخ مشخص فیلتر نمایید.

گزارش کلی کاربران تویت :

در این بخش گزارش های کلی از کاربران توییتر از قبیل پربازنشرترین کاربران ، محبوب ترین کاربر ، بیشترین منتشرکنندگان نشان داده می شود.در این نوع گزارش تصویر ، عنوان ، نام کاربری و تعداد نشان داده می شود در زیر گزارش یک فیلتر براساس تاریخ نیز قرار داده شده که شما می توانید خروجی را از یک تاریخ تا یک تاریخ مشخص فیلتر نمایید.

رترین مکان های انتشار توییت:

در این بخش گزارش های از مکان شامل اطلاعاتی از قبیل نام مکان وتعداد توییت های آن نشان داده می شود. در زیر گزارش یک فیلتر براساس تاریخ نیز قرار داده شده که شما می توانید خروجی را از یک تاریخ تا یک تاریخ مشخص فیلتر نمایید.

تاثیرگذارترین کاربران توییتر:

در این بخش گزارش های از کاربران شامل اطلاعاتی از قبیل نام کاربر وتعداد توییت های آن نشان داده می شود. در زیر گزارش یک فیلتر براساس تاریخ نیز قرار داده شده که شما می توانید خروجی را از یک تاریخ تا یک تاریخ مشخص فیلتر نمایید.

نمودارلحظه ای انتشار توییت

در این بخش نمودار لحظه ای انتشار توییت به صورت لحظه ای نشان داده می شود اطلاعات از قبیل زمان وتعداد تویت روی نمودار نشان داده می شود. در زیر گزارش یک فیلتر براساس تاریخ نیز قرار داده شده که شما می توانید خروجی را از یک تاریخ تا یک تاریخ مشخص فیلتر نمایید.

روند انتشارتوییت:

در این بخش نمودار روند انتشارتوییت به صورت میله ای نشان داده می شود اطلاعات از قبیل تاریخ و روند نشرتوییت روی نمودار نشان داده می شود. در زیر گزارش یک فیلتر براساس تاریخ نیز قرار داده شده که شما می توانید خروجی را از یک تاریخ تا یک تاریخ مشخص فیلتر نمایید.

# صفحه گزارشات تلگرام

ابرهشتگ های تلگرام ، ابرکلمات تلگرام:

در ابرهشتگ های تلگرام ، ابرکلمات تلگرام اطلاعاتی با ویژگی های از قبیل عنوان ابر ، عنوان کلمه مورد نظر و وزن کلمه نشان داده شده است در زیر هر ابرکلمات یک فیلتر براساس تاریخ نیز قرار داده شده که شما می توانید خروجی ابر کلمات را از یک تاریخ تا یک تاریخ مشخص فیلتر نمایید و یک دکمه لیست نیز در زیر ابر کلمات قرار داده شده که می توانید خروجی را به صورتی لیستی مشاهده نماید.

تعداد پست های منتشر شده و نرخ توزیع آن :

در این گزارش تعداد کاربران و نرخ توزیع پست ها به صورت کامل نشان داده می شود.

پربازدید ترین پست های تلگرام و آخرین پست های منتشر شده تلگرام :

در این بخش اطلاعاتی شامل عنوان پست ، تاریخ ، توضیح ، تعداد مشاهده پست نشان داده می شود. در زیر گزارش یک فیلتر براساس تاریخ نیز قرار داده شده که شما می توانید خروجی را از یک تاریخ تا یک تاریخ مشخص فیلتر نمایید.

پربازدیدترین کانال های تلگرام:

در این بخش گزارش های از شامل اطلاعاتی از قبیل نام کانال وتعداد مطلب آن نشان داده می شود. در زیر گزارش یک فیلتر براساس تاریخ نیز قرار داده شده که شما می توانید خروجی را از یک تاریخ تا یک تاریخ مشخص فیلتر نمایید.

پربازنشرترین کانال های تلگرام:

در این بخش گزارش های از شامل اطلاعاتی از قبیل نام کاربر و تعداد آن نشان داده می شود. در زیر گزارش یک فیلتر براساس تاریخ نیز قرار داده شده که شما می توانید خروجی را از یک تاریخ تا یک تاریخ مشخص فیلتر نمایید.

نمودار برترین منتشرکنندگان پست تلگرام

در این نوع گزارش عنوان ، نام کاربری و تعداد مشاهده به صورت نمودارنشان داده می شود در زیر گزارش یک فیلتر براساس تاریخ نیز قرار داده شده که شما می توانید خروجی را از یک تاریخ تا یک تاریخ مشخص فیلتر نمایید.

روندانتشارپست در تلگرام

در این نوع گزارش تاریخ و تعداد مشاهده به صورت نمودارنشان داده می شود در زیر گزارش یک فیلتر براساس تاریخ نیز قرار داده شده که شما می توانید خروجی را از یک تاریخ تا یک تاریخ مشخص فیلتر نمایید.

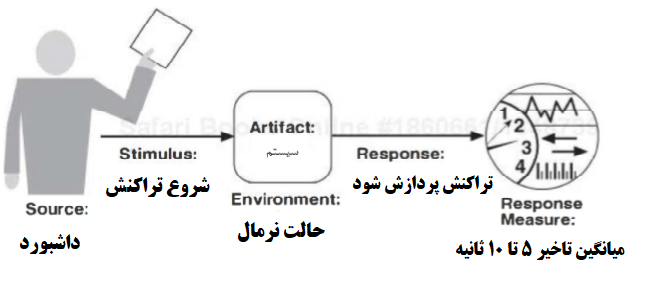
# فرم جستجو در سامانه

در این فرم برای کاربر جستجوی پیشرفته ای قرار داده شده است که شما می توانید عنوان خود را برای جستجو وارد نماییددر این فرم فیلتر های متعددی قرار داده شده است. که کاربر می تواند جستجوی خود را در بستر های مختلف در شبکه های اجتماعی و سایت های خبری با فیلتر های مختلف انجام دهد.

# سناریوهای غیروظیفه مندی:

1. کارایی :(performance)

داشبورد تراکنشی را تحت حالت نرمال آغاز می کند و داشبورد تراکنش را با میانگین تاخیر 5 تا 10 ثانیه پردازش می کند.



2.تاکتیک های برای افزایش کارایی:

* افزایش منابع: با توجه به بودجه شرکت استفاده از سرورهای با قدرت پردازش بالاتر و تعداد پردازشگر بالاتر و هم چنین استفاده از شبکه اینترنتی پرسرعت و افزایش تعداد کرال ها با GPU بیشتر به کاهش تاخیر کمک می کند.
* همروندی: استفاده از نخ ها برای پردازش درخواست ها باعث کاهش تاخیر می شود.
* (با توجه به چند نخی[[1]](#footnote-1) بودن این بخش، لیست لینک‏های استخراج شده بین چندین نخ (تعداد نخ‏ها قابل تنظیم است) تقسیم می‌شود و به صورت جداگانه هر نخ به واکشی اطلاعات از صفحه مرتبط با لینک می‌پردازد. برای مثال اگر 40000 هزار لینک خبری جمع‌آوری شده باشد و تعداد نخ‌ها را 4 در نظر گرفته باشید هر نخ 10 هزارتا از این لینک‌های خبری را گرفته و سپس هر نخ جداگانه و موازی با هم به واکشی اطلاعات از صفحات سایت‌های خبری می‌پردازند)
* زمتنبندی منابع: به دلیل رقابت بین سرویس گیرندها (client) برای به دست آوردن منابع سرویس دهنده (server) منبع باید طوری زمانبندی شود که تاخیر در انجام تراکنش ها را کاهش دهد.
* افزایش تعداد کراول ها با cpu پرقدرت تا صف کراول ها کمتر شده و با سرعت بیشتر و میزان کرال ها بیشتر شود.

### ایجاد glossary : تعریف برخی اصطلاحات

**بستر**: در اصطلاح، ما به هرکدام از شبکه‌های اجتماعی یک بستر می‌گوییم؛ از منظر این تعریف، توییتر یک بستر است همچنین تمام وب‏سایت‏های خبری اعم از خبرگزاری‏های و رسانه‏های منتشر کننده خبر را نیز بستر خبر می‏نامیم.

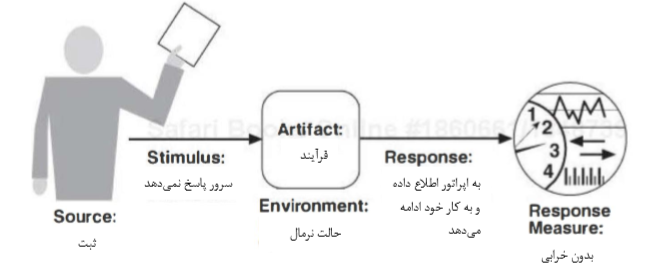
**گزارش**: به انواع تحلیل داده‌ها و ارائه نتایج آماری از یک یا چند بستر که داده‏های موجود را به انواع و اشکال نمایشی اعم از نمودارها، جداول، لیست‌ها و ... توصیف نماید.

**پارامتر**: به هر متغیری که دارای یک عنوان مشخص بوده و می‌تواند مقادیر مختلف بر اساس نوع تعریف شده اختیار کند. مثلا متغیر page در وب سرویس، یک پارامتر است که مقادیر عددی را می‌پذیرد.

* نگهداری چندین کپی از محاسبات : استفاده از چندین Rack برای نگهداری از کپی اطلاعات و هم چنین استفاده ار چندین سرور به جای یک سرور که که منجر به کاهش رقابت برای به دست آوردن یک منبع خاص می شود.
* استفاده از Load Balance قطعه نرم افزاری که کارهای جدید را به یکی از سرورهای در دسترس محول می کند.

دسترسی (Availability) :

داشبورد درخواست را از کاربر ثبت می کند و در صورت پاسخگو نبودن یکی از سرورها وضعیت را به در سیستم ثبت می کند و بدون از کار افتادن به کار خود ادامه می دهد و درخواست را ثبت می کند.



تاکتیک های برای افزایش دسترسی :

* Ping /echo : هنگامی که سامانه یا داشبورد به مشکل برخورد می کند میتوان ار پنگ اکو استفاده کنیم و ip سرو را پینگ کنیم تا در صورت قطع ارتباط یا خرابی در داشبورد آگاهی داشته باشیم و یا تاخیر رفت و برگشت را در نظر بگیریم با سرعت و ارتبلط ضعیف است و اگر آستانه زمان انتظار از حد معینی بیشتر شود سرور را از کار افتاده در نظر میگیریم.
* Exception handling : زمانی که سرور پاسخ نمی دهد با این روش از کار افتادن سیستم جلوگیری کرده و جریان نرمال سیستم را تغییر می دهیم و سعی می کنیم به سرور دیگه ای وصل شویم.
* Rerry : تلاش مجدد – این امکان وجود دارد تا در خواست ارتباط با سرور دوباره ارسال شود زیرا ممکن است سرور از کار نیافته باشد و تنها یک قطعی موقت در ارتباط با شبکه باشد.
* تراکنش(Transcation): با نوشتن عملیات ثبت درخواست درون یک تراکنش از انجام شدن ناقص آن جلوگیری می کنیم و در صورت برقرار نشدن ارتباط با سرور کل عملیات را برمیگردانیم.

### قالب کلی درخواست‌های[[2]](#footnote-2) وب سرویس

هر درخواستی که به سمت وب سرویس ارسال می‏شود باید از یک قالب کلی پیروی نماید در غیر اینصورت، نتایج نامعتبر و یا با خطا روبرو خواهد شد. درخواست‏های معتبر از سمت کاربر باید حداقل ۴ مورد اول از موارد زیر را دارا باشند. با داشتن تنها ۴ مورد اول، نتایج گزارش موردنظر به صورت پیشفرض (بدون اعمال شروط جستجوی پیشرفته) نمایش داده می‌شود.

* قسمت ثابت درخواست
* نام بستر
* ثابت Query
* عنوان گزارش (Class)
* پارامتر صفحه بندی[[3]](#footnote-3) (page)
* پارامترهای جستجوی پیشرفته
* پارامتر مقایسه در یک گزارش (comp)

### مشاهده تمام گزارش‏ها موجود در وب‏سرویس

برای مشاهده تمام گزارش‏ها موجود به همراه لینک درخواست و عنوان گزارش می‌توان از بخش help وب سرویس استفاده نمود. در بخش help پس از انتخاب بستر مورد نظر، درخواستی به سمت وب سرویس ارسال می‌شود که از این طریق می‌توان تمام گزارش‏ها آن بستر را مشاهده نمود. برای دسترسی به این بخش، در لینک درخواست‏های وب‏سرویس به جای query باید از کلمه help استفاده شود و در ادامه نام بستر ذکر گردد.

1. MultiThread [↑](#footnote-ref-1)
2. Requests [↑](#footnote-ref-2)
3. Pagination [↑](#footnote-ref-3)