

استفاده از یک ابزار آماده برای تبدیل صوت فیلم به متن (انگلیسی)

تهیه و تنظیم: مبین خیبری

شماره دانشجویی: 994421017

چکیده:

در سطرهای پیش رو قصد داریم متن گفت‌وگوها و یا محتوای دیالوگ‌های یک فیلم را با استفاده از موتورهای هوش مصنوعی توسعه داده شده استخراج کرده و نتیجه‌ی حاصل را با محتوای اصلی گفت‌وگو میان شخصیت‌ها مقایسه کنیم. برای انجام این کار ابتدا نیاز به یک فیلم کوتاه داریم که شامل مقادیری مناسبی از گفت‌وگوها باشد. بخش تحلیلی این مسئله بر اساس متون انگلیسی اجرا خواهد شد، بنابراین پیش‌فرض در اینجا این است که فیلم مذکور، اصالتاً به زبان انگلیسی ساخته و زیرنویس اصلی آن هم به همین زبان نگاشته شده.

برای انجام این سری تحلیل‌ها ابتدا باید اقدامات لازم را در سه مرحله‌ی مجزا انجام دهیم:

مرحله‌ی نخست: ابتدا باید به وسیله‌ی برخی ابزارها بخش صوتی مربوط به فیلم را استخراج کنیم.

مرحله‌ی دوم: در این قسمت باید به کمک ابزارهای توسعه داده شده به وسیله‌ی هوش مصنوعی، فایل صوتی استخراج شده در مرحله‌ی گذشته را به متن تبدیل کنیم.

مرحله‌ی سوم: در نهایت باید به کمک یک ابزار تفاوت‌سنج، میزان تفاوت و اختلاف میان زیرنویس اولیه و زیرنویس تولید شده توسط هوش مصنوعی را بررسی کنیم.

در ادامه به تفصیل درباره‌ی این مراحل توضیح خواهیم داد:

مرحله‌ی نخست:

اولین مرحله‌ی تحلیل و بررسی این مسئله را با انتخاب یک فیلم کوتاه آغاز می‌کنیم.

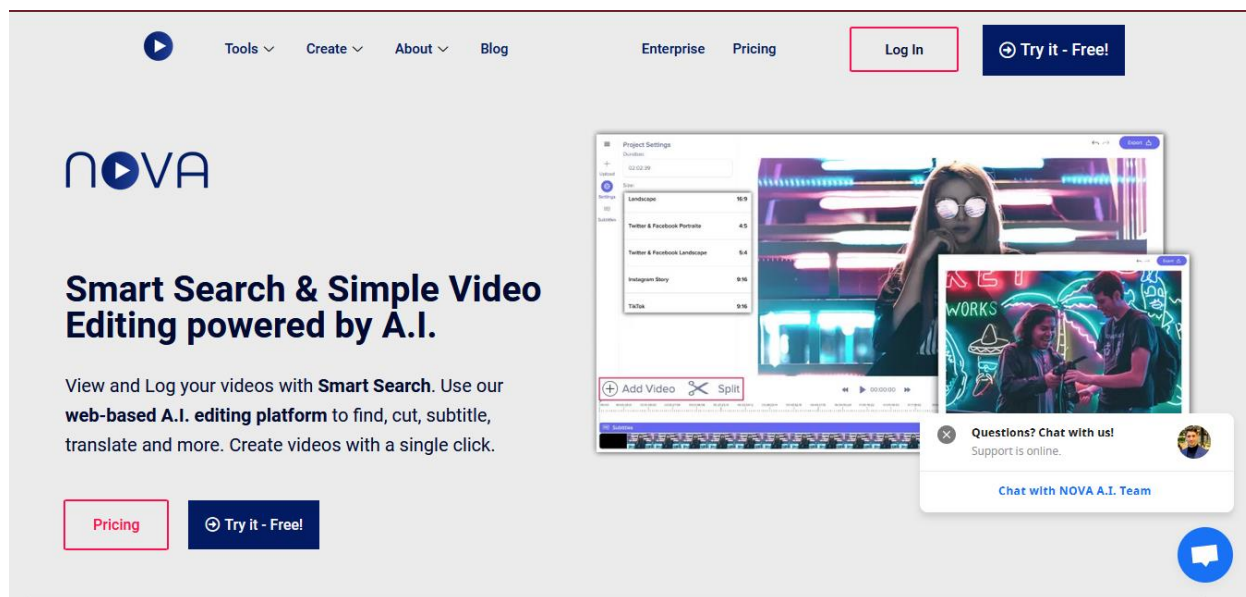
فیلم انتخابی "بهترین دوست" نام دارد که در سال 2020 توسط "باهارات داریرا" ساخته شده. این فیلم کوتاه که در ژانر سینمایی کمدی سیاه قرار می‌گیرد، به زبان انگلیسی و در سینمای هند ساخته شده. پس از کسب موفقیت‌هایی در جشنواره‌های بین‌المللی فیلم کوتاه، سازندگان این فیلم نسخه‌ی کامل آن را به‌طور رایگان بر روی وبسایت یوتیوب آپلود کردند.

نسخه‌ی کامل این فیلم 8 دقیقه‌ای برای آغاز روند نگاه تحلیلی به تأثیرات هوش مصنوعی بر صنعت سرگرمی، انتخاب هوشمندانه‌ای به نظر می‌رسد.

سپس نیاز داریم که فایل صوتی حاوی گفت‌وگوهای انجام شده درون فیلم را از درون آن استخراج کنیم.

ابزارهای بسیار زیادی برای استخراج صوت از فایل‌های ویدیویی توسعه داده شده‌اند. این ابزارها قابلیت اجرا بر روی سیستم‌های خانگی و نیز سرورهای ابری را نیز پشتیبانی می‌کنند.

انتخاب ما در این قسمت استفاده از ابزار Nova است. در سطرهای زیر به‌طور مختصر به قابلیت‌های این هوش مصنوعی که در واقع یک ادیتور آنلاین قدرتمند است، خواهیم پرداخت. برای وفاداری به مرجع اصلی، این توضیحات به زبان انگلیسی آورده شده‌اند.



درباره‌ی هوش مصنوعی Nova:

Nova A.I. is an end-to-end video editing solution supporting you in every step of the video editing and management process. From the initial steps of identifying the perfect scene and finding the right clips to auto subtitling, translation and editing.

Nova auto analyses your videos and labels everything that happens in them. Enabling you to instantly deep-search your video, audio, and archival libraries for specific moments categorized by emotions, objects, people, actions, scene types, words and more. Making video content and archives 100% searchable.

The software automatically transcribes, translates, and subtitles 40+ languages and allows users to generate multiple new clips from original content in seconds.

What is Nova A.I.?

Nova is a must have tool in every video professional arsenal. It's a new generation automatised video labelling and archiving software powered by solid computer vision search engine.

Why should I choose Nova A.I.?

If you want to unlock the full power of your video content, discover hidden data that can be leveraged and used to surf your entire content library quickly, Nova is the right choice for you.

Which file formats does Nova A.I. support?

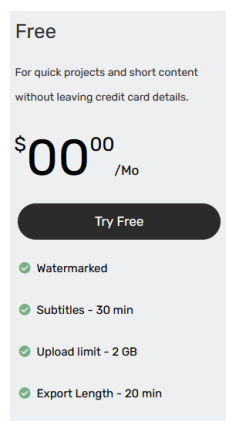
Nova A.I. is compatible with most video and audio formats. These include: MP3, WAV, PNG, JPEG, MOV, MP4, RAW, MXF, MPEG and more. Nova supports all major media file types. If you have files that are not supported by Nova, do not hesitate to drop us a message via the chat and we will do our best to sort that out for your.

Can I edit videos with Nova A.I.?

At the moment Nova A.I. is not a standalone software. The main function of the platform is turning your content into searchable digital assets, but some video editing functions will be available with future updates.

How does Nova A.I. work?

Nova runs on many various Artificial Intelligence technologies provided by aws media intelligence and put together by our genius tech warriors from ukraine!



برای دسترسی به ویدیوی "بهترین دوست" می‌توانید از لینک زیر استفاده کنید:

https://www.youtube.com/watch?v=NpzNC_35RR0

برای دسترسی به زیرنویس اصلی این فیلم می‌توانیم از ابزارهای آنلاین بسیار متنوعی استفاده کنیم. انتخاب ما در این قسمت استفاده از DownSub است. این سرویس اینترنتی به کاربران این امکان را می‌دهد که زیرنویس‌های افزوده شده به ویدیوهای وبسایت یوتیوب را به صورت رایگان و با فرمت‌های مختلف متنی دانلود کنند.

نمای اصلی وبسایت مذکور بدین شکل است:

The screenshot shows the DownSub website. At the top is a navigation bar with links: HOME, HISTORY, SUPPORTED SITES, CONTACT, DONATE, and LANGUAGE. Below the navigation bar is a search bar with the placeholder text "Enter a link to download subtitles. Ex: https://youtu.be/rN7yhDI1cuk" and a green "DOWNLOAD" button. The main content area is divided into two columns. The left column is titled "Download Subtitles" and contains a paragraph explaining that DownSub is a free web application that can download subtitles directly from YouTube, ViU, Viki, and more. It also lists supported subtitle formats: SRT, TXT, VTT. The right column is titled "Recent Updates" and lists several updates: MGTV, Erosnow, Brightcove, and Weverse, each with an "ADDED" status. Below these columns is a "How to use" section with a "Copy-paste" icon and a "Tricks" icon. The "Copy-paste" section lists three steps: 1. Paste the link of video you want to download subtitles. 2. Click Download button to begin extracting subtitles. 3. Select the subtitles format and language you want to download, then click Download button. The "Tricks" section shows two examples of how to use the service: "Add subtitle.to/ before the URL and press Enter" and "Or add subtitle.to/ between the URL and press Enter".

بعد از دانلود زیرنویس‌های اصلی فیلم به سراغ وبسایت هوش مصنوعی یادشده رفته و به کمک موتورهای قدرتمند آن، ویدیو را به فایل صوتی با فرمت دلخواه تبدیل می‌کنیم.

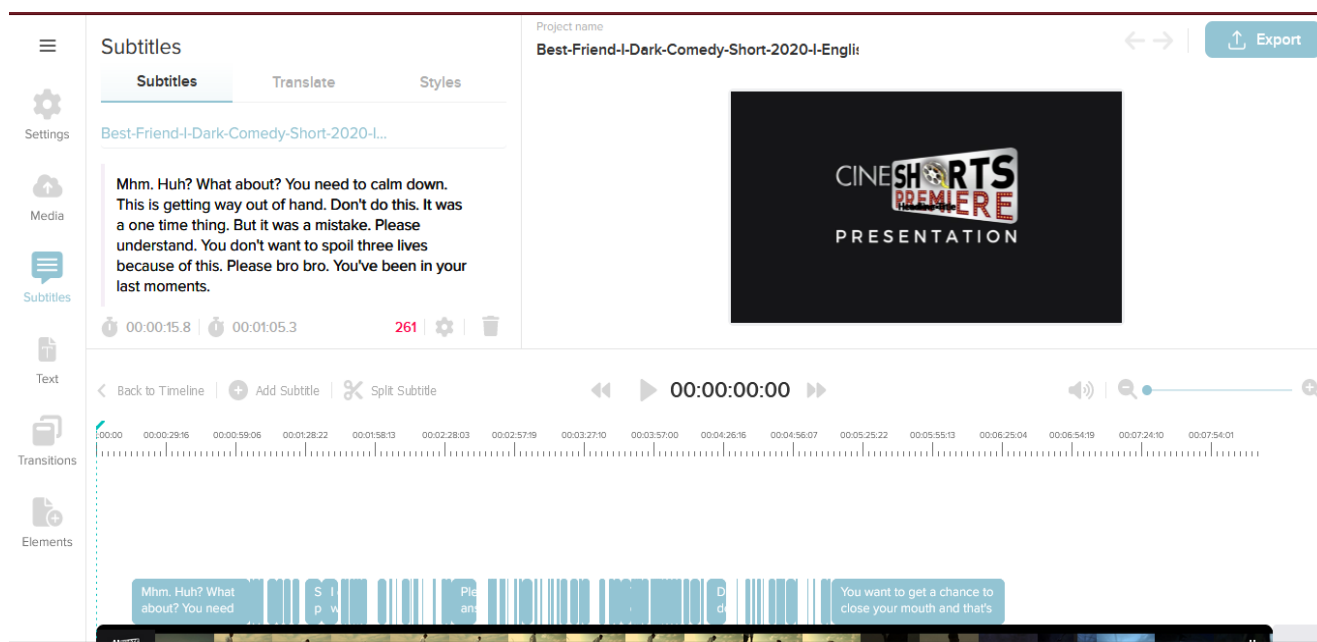
The screenshot shows a video editing software interface. On the left is a sidebar with icons for Settings, Media, Subtitles, Text, Transitions, and Elements. The main area is divided into two sections. The top section is titled "Project Settings" and contains fields for "Duration" (00:08:20) and "Size" (Original). The bottom section is a video player with a timeline. The timeline has a red bar at the beginning and a blue bar at the end. The video player has a play button and a progress bar. The project name is "Best-Friend-I-Dark-Comedy-Short-2020-I-English-Shor".

مرحله‌ی دوم:

در این مرحله نیاز داریم که با استفاده از یکی از موتورهای هوش مصنوعی‌های توسعه داده شده و یا ابزارها و زبان‌های برنامه‌نویسی مختلف، برای فایل صوتی استخراج‌شده، زیرنویس مصنوعی تولید کنیم.

خوشبختانه ابزار Nova این قابلیت را هم شامل می‌شود. ضمن اینکه می‌توان زیرنویس تولید شده را توسط همین ابزار به فرمت‌های مختلف متنی استخراج کرد.

نمونه‌ای از قابلیت‌های آن را در شکل زیر مشاهده می‌کنید:



مرحله‌ی نهایی:

در این بخش با استفاده از فایل‌های مربوط به زیرنویس‌ها، و نیز به کمک ابزاری به نام CopyLeaks استفاده می‌کنیم.

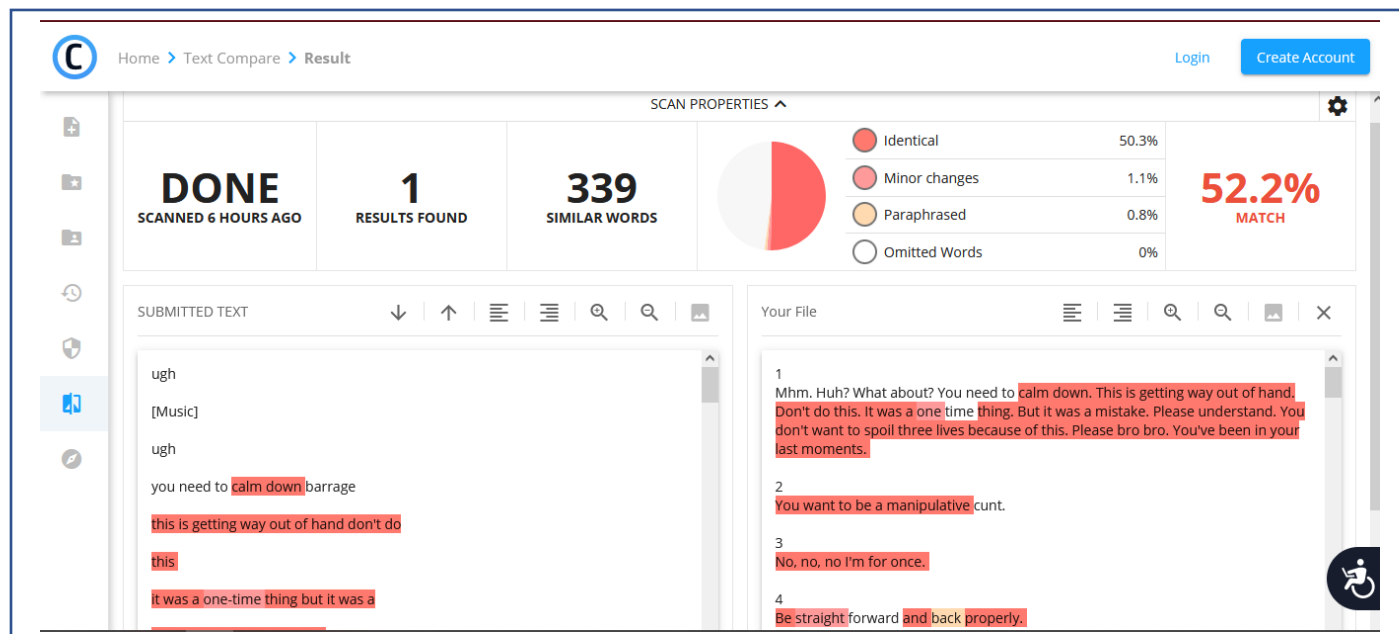
این ابزار به ما کمک خواهد کرد تا مرحله‌ی مشابهت‌سنجی میان متون را به راحتی و با دقتی نسبتاً بالا به سرانجام برسانیم.

نتایج تحلیل انجام‌شده در صفحات بعد آورده شده‌اند. بر اساس برآوردهای این ابزار، 52 درصد مشابهت میان دو زیرنویس وجود دارد.

ذکر این نکته لازم به نظر می‌رسد که لهجه‌ی هندی بازیگران فیلم، سانسور برخی عبارات به کارگرفته‌شده در آن توسط سیستم اتوماتیک تولید زیرنویس یوتیوب و عامیانه بودن بیش از حد برخی جملات، سختی کار را برای موتور هوش مصنوعی دوجندان کرده.

ما نیز برای سنجش میزان موفقیت هوش مصنوعی، عمدا دخالتی در جهت آسان‌تر کردن فرآیند انجام ندادیم.

نتایج حاصل از این تحلیل، در شکل زیر به‌طور مفصل آورده شده:



منابع استفاده‌شده برای تهیه‌ی این گزارش:

- I. https://www.youtube.com/watch?v=NpzNC_35RR0
- II. <https://vidiget.com/home13>
- III. <https://downsub.com/>
- IV. <https://text-compare.com/>
- V. <https://app.copyleaks.com/>
- VI. <https://app.wearenova.ai>