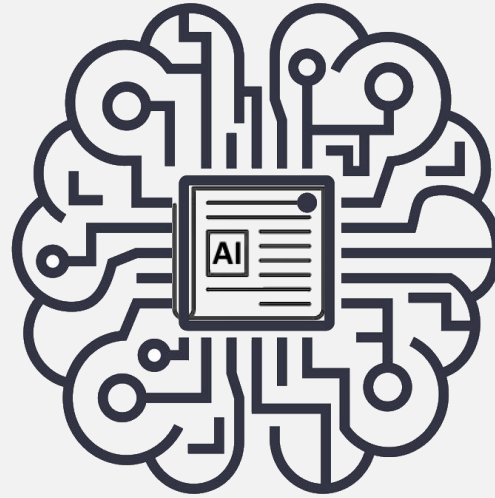


Presentation on Final Project of Artificial Intelligence (CSCI-6660-01-S20)



Automated Podcast Generation

Submitted to:

Dr. Vahid Behzadan
Assistant Professor of Computer Science and Data Science
University of New Haven

Submitted by:

Bibek Upadhayay
bupad1@unh.Newhaven.edu

Problem

- More than 51% of US population has listened to Podcast in 2019
- Only 35% of those listen to the whole podcast
- The reason behind this is that the content was not interesting for them

Q. How can we make Interesting Podcast?

Let's use **AI** to create Podcast.



Objective

- To create a project to satisfy the requirement of final project for Artificial Intelligence (CSCI-6660-01-S20)
- To use the available knowledge of artificial intelligence and technology to create an application
- To develop an automated system that creates a podcast by itself and publish the podcast to respective platform
- To develop an automated podcast generation system that automatically creates the podcast by using available content in the web

Approach



- On the left it is typical podcast system, where podcast host creates and publish the podcast on respective platform

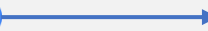
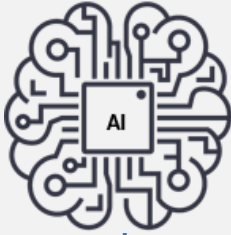
Approach

- Our objective is to replace the podcast host by generating the podcast using AI
- We are going to use AI to create an interesting content for a particular user
- We will be creating a podcast for an artificial intelligence news
- The technology we are going to use are
 - Python
 - Google WaveNet for speech generation
 - Google API
 - Apple iTunes platform to host the podcast

Approach



Website #1



System Algorithm



Podcast Episode



HOW DOES THE SYSTEM WORK?

Step1: We will extract the news content from the respective website

Step2: We will filter the content and separate each news.

Step3: the system will takes those Speech Files and use Google API to convert them into speech files

Step 4: the system will combine the speech files and background music to create a full podcast episode

Step 5: The system will generate and update the .XML file and submit to Apple iTunes to publish the podcast

Evaluation Methodology

- An executable python code that will do the following things
 - Display the title and content of the cleaned news
 - Will create an audio podcast generated by the system
 - An .XML file that auto-updates for publishing podcast
 - Link to the iTunes for published and approved podcast.

Methodology

- Environment Setup
 - Create a virtual environment
 - `pip install google-cloud-texttospeech`
 - `pip install feedparser`
 - `pip install AudioSegment`
 - `pip install pydub`
- Google Credentials for API: Get the .json credentials from Google Cloud
- .XML File setup: Need to follow the iTunes podcast rss format for the .XML file
- The program extract the text content from <https://artificialintelligence-news.com/feed/> . The text is first cleaned via `cleanText()` function

Few important notes:

- Google only allows <5000 characters to be converted so the recursive function `feed_fragmentation()` function slice the larger text and convert them into smaller audios and later these audios are combined.
- For each episode the xml file need to be updated
- For the headline audios, the algorithm makes sure that the background music automatically increase-decrease in volume.

For every episode, this code is added in .XML file

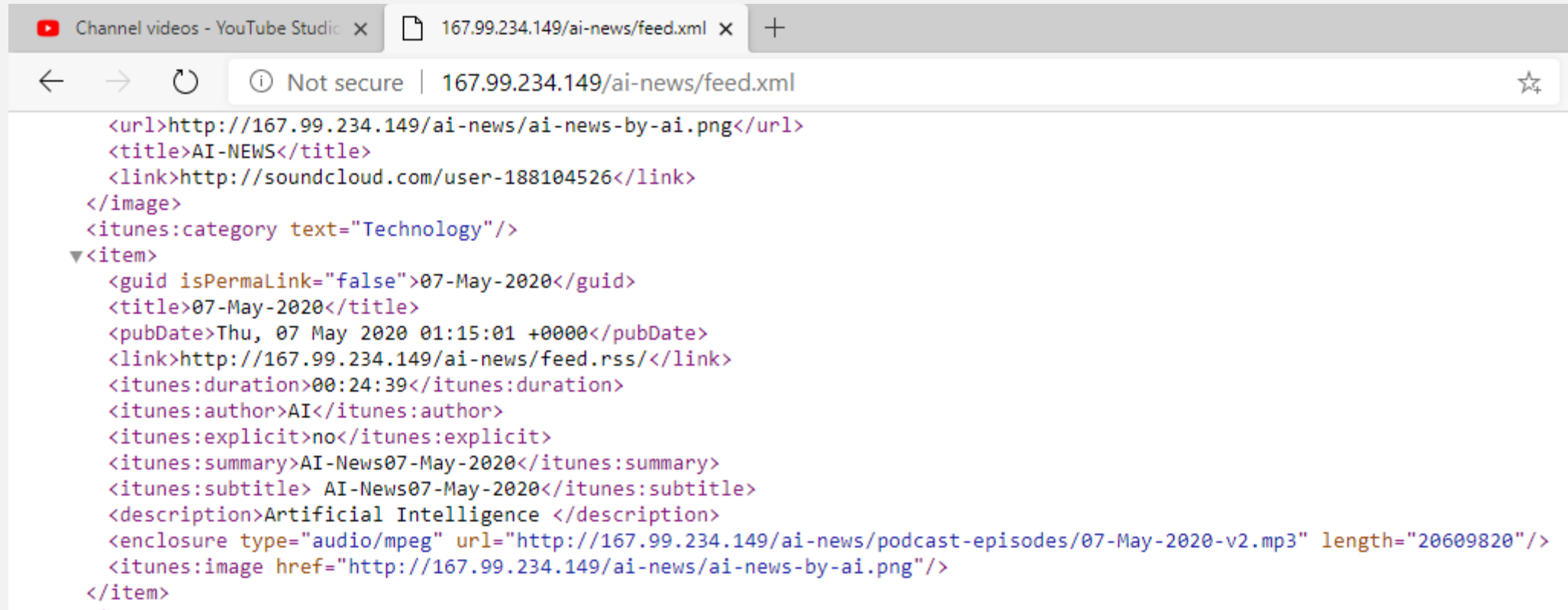
```
<item>
<guid isPermaLink="false">""+this_day+""</guid>
<title>"" + Title +""</title>
<pubDate>""+PubDate+""</pubDate>
<link>""+Link+""</link>
<itunes:duration>""+Duration+""</itunes:duration>
<itunes:author>AI</itunes:author>
<itunes:explicit>no</itunes:explicit>
<itunes:summary>""+Summary+""</itunes:summary>
<itunes:subtitle> ""+Subtitle+""</itunes:subtitle>
<description>""+Description+""</description>
<enclosure type="audio/mpeg" url= "http://167.99.234.149/ai-news/podcast-episodes/"" \
    |+todays_date+""-v2.mp3" \ length="20609820"/>
<itunes:image href="http://167.99.234.149/ai-news/ai-news-by-ai.png"/>
</item>

"" ""
```

Files generated after the running the code.

```
root@AIPodcast: /home/ai2/abc
root@AIPodcast:/home/ai2/abc# ls
Big_Day.mp3  bin          head5.mp3  news-music.mp3  news_3_1.mp3  save_time.py
__pycache__  date.txt     head6.mp3  news_0_1.mp3    news_3_2.mp3  share
ai_news.py   final_1.mp3 include     news_0_2.mp3    news_4_1.mp3  try.py
ainews.py    head1.mp3   key.json   news_1_1.mp3    news_4_2.mp3
ainews2.py   head2.mp3   kl.py      news_1_2.mp3    news_5_1.mp3
ainews3.py   head3.mp3   lib        news_2_1.mp3    news_5_2.mp3
audio_files  head4.mp3   lib64      news_2_2.mp3    pyenv.cfg
root@AIPodcast:/home/ai2/abc#
```

.XML files updated

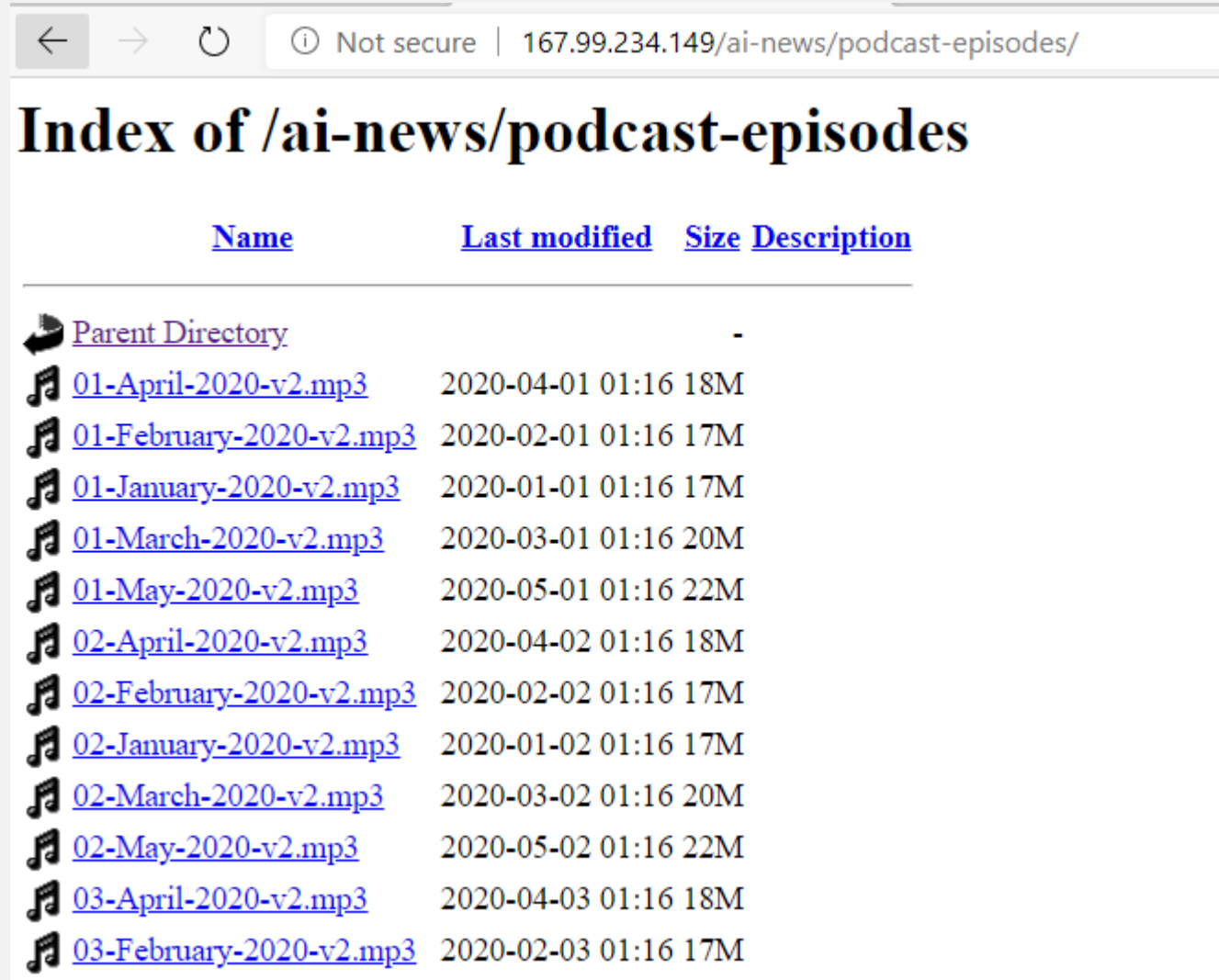


A screenshot of a web browser window. The address bar shows the URL `167.99.234.149/ai-news/feed.xml` with a status of "Not secure". The page content displays an XML document. The visible XML tags include `<url>`, `<title>`, `<link>`, `</image>`, `<itunes:category text="Technology"/>`, `<item>`, `<guid isPermaLink="false">`, `<title>`, `<pubDate>`, `<link>`, `<itunes:duration>`, `<itunes:author>`, `<itunes:explicit>`, `<itunes:summary>`, `<itunes:subtitle>`, `<description>`, `<enclosure type="audio/mpeg" url="http://167.99.234.149/ai-news/podcast-episodes/07-May-2020-v2.mp3" length="20609820"/>`, `<itunes:image href="http://167.99.234.149/ai-news/ai-news-by-ai.png"/>`, and `</item>`.














```
<url>http://167.99.234.149/ai-news/ai-news-by-ai.png</url>
<title>AI-NEWS</title>
<link>http://soundcloud.com/user-188104526</link>
</image>
<itunes:category text="Technology"/>
<item>
  <guid isPermaLink="false">07-May-2020</guid>
  <title>07-May-2020</title>
  <pubDate>Thu, 07 May 2020 01:15:01 +0000</pubDate>
  <link>http://167.99.234.149/ai-news/feed.rss</link>
  <itunes:duration>00:24:39</itunes:duration>
  <itunes:author>AI</itunes:author>
  <itunes:explicit>no</itunes:explicit>
  <itunes:summary>AI-News07-May-2020</itunes:summary>
  <itunes:subtitle> AI-News07-May-2020</itunes:subtitle>
  <description>Artificial Intelligence </description>
  <enclosure type="audio/mpeg" url="http://167.99.234.149/ai-news/podcast-episodes/07-May-2020-v2.mp3" length="20609820"/>
  <itunes:image href="http://167.99.234.149/ai-news/ai-news-by-ai.png"/>
</item>
```

<http://167.99.234.149/ai-news/feed.xml>

Output files generated in the server



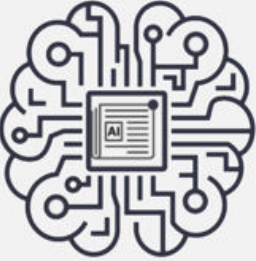
The screenshot shows a web browser window with the address bar displaying "167.99.234.149/ai-news/podcast-episodes/". The page title is "Index of /ai-news/podcast-episodes". Below the title is a table with four columns: "Name", "Last modified", "Size", and "Description". The table lists a "Parent Directory" and 15 MP3 audio files, each with a date, time, and size. The files are organized by month and day, with some files having a version number (v2).

Name	Last modified	Size	Description
 Parent Directory		-	
 01-April-2020-v2.mp3	2020-04-01 01:16	18M	
 01-February-2020-v2.mp3	2020-02-01 01:16	17M	
 01-January-2020-v2.mp3	2020-01-01 01:16	17M	
 01-March-2020-v2.mp3	2020-03-01 01:16	20M	
 01-May-2020-v2.mp3	2020-05-01 01:16	22M	
 02-April-2020-v2.mp3	2020-04-02 01:16	18M	
 02-February-2020-v2.mp3	2020-02-02 01:16	17M	
 02-January-2020-v2.mp3	2020-01-02 01:16	17M	
 02-March-2020-v2.mp3	2020-03-02 01:16	20M	
 02-May-2020-v2.mp3	2020-05-02 01:16	22M	
 03-April-2020-v2.mp3	2020-04-03 01:16	18M	
 03-February-2020-v2.mp3	2020-02-03 01:16	17M	

<http://167.99.234.149/ai-news/podcast-episodes/>

iTunes Podcast channel

Apple Podcasts Preview



AI-NEWS-BY-AI

150 episodes

AI-NEWS podcast is created for AI enthusiasts. For any person it is a hassle to go to different websites and search for AI news and updates. Now the users can listen to this daily podcast and can get all the updates and news around the world on AI.

AI NEWS by AI
AI-NEWS-BY-AI, University of New Haven
Technology

[Listen on Apple Podcasts ↗](#)

MAY 6, 2020
07-May-2020 >
AI-News07-May-2020
[▶ PLAY](#) 24 min

MAY 5, 2020
06-May-2020 >
AI-News06-May-2020
[▶ PLAY](#) 24 min

MAY 4, 2020
05-May-2020 >
AI-News05-May-2020
[▶ PLAY](#) 24 min

<https://podcasts.apple.com/us/podcast/ai-news-daily/id1491115711>

Thank you!