



Master of  
Management Analytics  
*Toronto*

**Course Number: MMA 823**  
**Course Name: Big Data Analytics**

**Assignment Name: Individual Assignment**  
**Due Date: Jan 29, 2025 11:59 pm**

**Team Name: Team Gordon**

Student Name	Student Number
Anthony Ramelo	20499391

Order of files:

Filename	Pages	Comments and/or Instructions
MMA 890 Individual Assignment 2	4	

**Additional Comments:**

--

## **Topic: Summarization of Meetings Using Natural Language Processing in Highly Regulated Sectors**

### **Section 1:**

In big corporations, especially in sectors like financial services and law firms, employees handle sensitive information that needs to be handled quickly and accurately. These action items are collected through meetings, where employees spend a lot of time taking meeting minutes and making sure that all important items are captured. However, during long meetings, some key points and action items are occasionally missed due to a lack of focus.

With advancements in Natural Language Processing (NLP) tools, employees can complete meeting minutes much faster and more accurately. Tools like Athena GPT, Co-Pilot, and ChatGPT can transcribe conversations in real-time and generate summaries and action items for all meeting participants. This not only saves time but also ensures accuracy, as action items are clear and assigned accurately.

### **Section 2:**

We suggest using NLP tools like Athena GPT, Co-Pilot, and ChatGPT to transcribe meeting audio into text, reducing errors saving time, and making sure we have regulatory guidelines for highly regulated sectors. This could be done by integrating a Retrieval Augmented Generation (RAG) system, this system learns from organizational documents to ensure summaries are accurate. NLP tools would then analyze the transcript to extract key items. An employee would then review the summary and action items produced by the NLP to ensure accuracy before sending the items out. This process would help turn audio/transcripts into actionable items while following regulatory guidelines.

### **Section 3:**

#### **Time Savings:**

- In a 60-minute meeting, an average person has about 7800 words (VirtualSpeech, n.d.). This is about 7800 words per 60-minute meeting. Manually transcribing and summarising a meeting take about 3-4 hours per meeting (SpeakWrite, n.d.)
- By using NLP tools like Athena GPT, Co-Pilot, and ChatGPT will take about 5 minutes to complete the meeting summary and action items for people to complete. This is a 97.92% reduction in the amount of time an employee spends on creating meeting minutes per meeting.

#### **Cost Savings:**

- If an employee attends two meetings per day and has an hourly wage of \$20, this is going to save \$1200 per employee every week.

#### **Section 4:**

In highly regulated sectors such as financial services and law firms, it is important to ensure that an NLP model is developed in a secured environment. Below are a few items to consider:

1. Developing the model in a secure environment to ensure that no privileged or sensitive information is inadvertently used for training or future applications.
2. Train the model to recognize and understand the terminology specific to the department or division of the employee using the model. This improves the model's ability to accurately classify and interpret industry terminology.
3. Launching the model in a low-impact division or department initially. The approach allows testing the model's performance with minimal consequences. Any mistake made will have a low cost of impact.
4. Monitoring the output of the NLP model closely. If the model meets expectations, it can be expanded to another low-impacting group to minimize risk while scaling.

We would be able to collect data from past transcripts of already recorded meetings. We would also utilize documentation about the department, including notes and other relevant materials, to ensure accuracy.

Below are a couple of requirements for the process:

1. We would require an NLP tool that is able to turn audio accurately into a transcript.
2. A repository of common terminology used by the department/ division, which the RAG system could learn from.

#### **What type of modeling is required.**

We want to use a Retrieval Augmented Generation (RAG) system, this system learns from organizational documents to ensure summaries are accurate. As well as use an NLP tool that will be able to summarise the meeting.

We would test the model by comparing the model summary output to the manually generated summary. Making sure that all terminology was captured correctly and the employees assigned to the action items were also identified. If the model did not perform correctly, we would then retrain the model with the user feedback.

Reference:

VirtualSpeech. (n.d.). Average speaking rate and words per minute. Retrieved January 22, 2025, from <https://virtualspeech.com/blog/average-speaking-rate-words-per-minute>

SpeakWrite. (n.d.). How to transcribe meeting minutes. Retrieved January 22, 2025, from <https://speakwrite.com/blog/how-to-transcribe-meeting-minutes/>