Project Peggy Olsen

DATA DRIVEN CREATIVITY

Copenhagen, Denmark

GroupM Mentor: Kristjan Brødeskrift

Capstone Project By

Jixin WANG & Melody FONG

MIT Advisor: Dean Eckles

MIT Mentor: Leon Yao

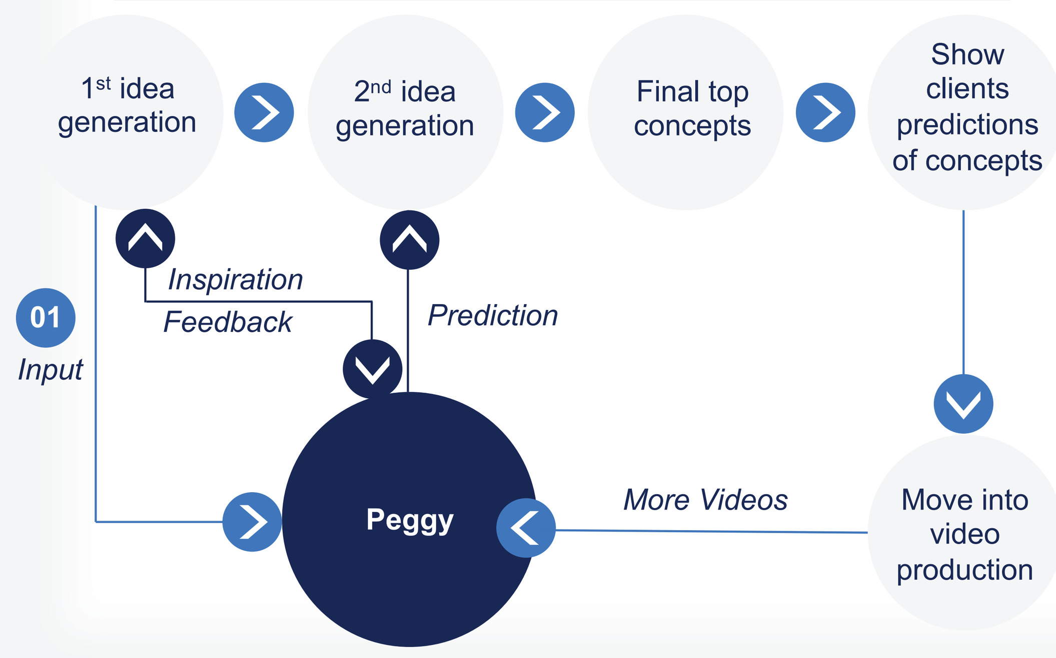
Why Project Peggy Olsen

GroupM is a leading global Media Agency. Video creation is one of its core business lines. Video production is costly and hence GroupM wants to make sure the idea is right before putting it into production. The creative team spends tons of time searching for relevant videos to gain inspirations and refine their proposal. Project Peggy Olsen is designed to be a data-driven solution to facilitate idea generation. This project is an crucial strategic move for GroupM to be the industry pioneer of fueling creativity with data.

What is Project Peggy Olsen

Project Peggy Olsen is a powerful recommendation engine that takes in mood board, a series of pictures uploaded by creative team conveying their video ideas, and outputs similar videos to provide inspirations.

Where Peggy Olsen Lies in the Creative Process



Why We Name it after Peggy Olsen

Peggy Olson a fictional character in Mad Men. Peggy symbolizes the disruption of a male dominated marketing industry by bringing the female perspective to creativity. She represents the change that leapfrogged an industry to the next level. What Peggy was for 1960s is what data science will be for the future. Hence, we name our recommendation engine after "Peggy Olson".

Vision of Project Peggy Olsen

- Multidimensional Feature Sets to Ensure Holistic Comparison
- Easy-to-use User Interface for the Creative Team to Adopt in Video Creation
- Machine Learning Model to Identity Similar Videos
- Industry Leader for Data Drive Creativity

Project Timeline

- Define Project Objective and Scope — February
- March — Feature Discovery and Extraction
- Define Project Objective and Scope — April
- May — Feature Study
- Experiment with Different Methods — June
- July — Decide on Final Methodology
- Prepare Outputs and Presentations — August

How Peggy Olsen Works

Map Input Data	Distance Based on Labels	Distance Based on Other Features	Weight Total Distance	Filter Options	Recommendations
Map uploaded mood borad pictures into the same feature space as existing video ads	Use Jaccard Distance to calculate the distance between mood board data and indiviudal video data based in terms of content	Standardize all features besides labels and calculate the euclidean distance between mood board data and individual video data	Combine two different distances together while assigning higher weight to label feature	Filter option for the creative team to filter on some specific features. For example, they can filter videos by industry or Video length	Recommend most relevant or similar videos by their similarities with the moodboard data

Feature Study & Selection

Data

- 6,369 historical video ads stored in GroupM cloud platform
- Google API is utilized to extract both visual and audio features
- 62,294 scene-level entries and 17 variables, describing each video ads from various perspectives

Color Features

Hue
Warm, Cold, Neutral

Saturation
Vivid vs Greyish

Lightness
Brightness of the scene

Emotions Detected in Videos

To figure out the best way to aggregate scene-level emotions into video-level, we tested on 5 different techniques:

- Aggregated Emotion $= \sum_{k=0}^{n} Emotions^k$
- Aggregated Emotion $= \frac{1}{n}\sum_{k=0}^{n} Emotions^k$
- Aggregated Emotion $= Max(Emotions^k)$
- Aggregated Emotion $= \begin{cases} 1 & \text{If the emotion exists} \\ 0 & \text{Otherwise} \end{cases}$
- only take the dominant emotion for each video

After inspecting the emotions in sample videos, we chose to aggregate emotional feature by taking the maximum value within each video

Labels

One-Hot Encoding

We formulate a table which looks like below.

	Tree	Car	Human	Sky	Indoor	Text	......
Video 1	0	1	1	1	0	0	......
Video 2	0	0	0	0	0	1	......
Video 3	1	0	0	1	0	0	......
......							......

We input 1 if the commercial (row) contains the label (column) and 0 otherwise. We then compute the Jaccard distance between the target selected sample videos and the other videos and inspect them to see how similar they are in terms of commercial content.

$$J(A,B) = \frac{|A \cap B|}{|A \cup B|} = \frac{|A \cap B|}{|A| + |B| - |A \cap B|}$$

$$D_J(A,B) = 1 - J(A,B) = \frac{|A \cup B| - |A \cap B|}{|A \cup B|}$$

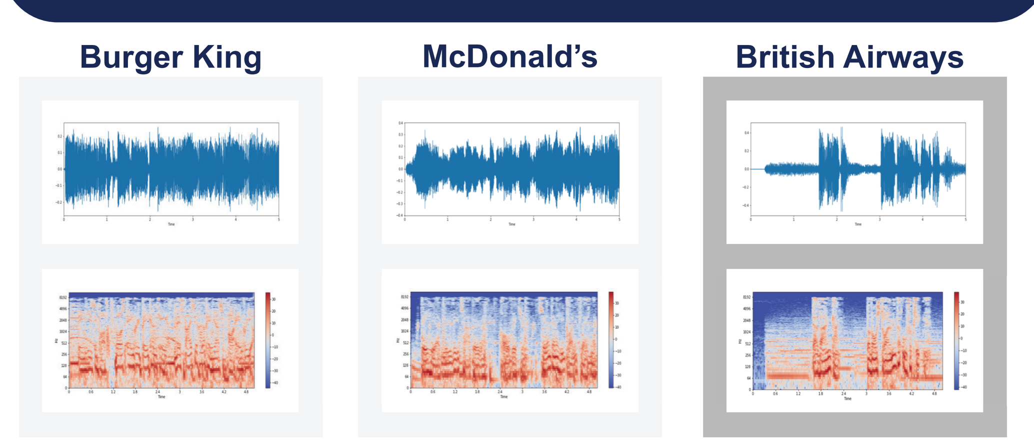
Word2Vec

- Train Word2Vec models using different corpuses: video labels, tweet subset, news, Wikipedia, Q&A forum
- Evaluate Word2Vec models using current dataset: intra- and inter-group distance
- Adopt top 2 performing Word2Vec model to find most similar videos and conduct internal blind test

Decision for Final Model

Internal blind test results show that binary one-hot encoding finds more relevant videos 100% of the time. After exploring multiple encoding methods, we are more comfortable with choosing binary one-hot encoding.

Audio Features



Audio tracks are embedded as the spectrograms, a heat map describing the density of frequency. We then extract 12+ features from the spectrograms and aggregate them into video-level. Feature examples: centroid, contrast, roll-off, bandwidth, flatness, mel-frequency cepstral coefficients, zero crossing rate, etc…

Stand Alone Performance Summary

- Conduct internal blind tests with 34 videos in the test set
- Naïve baseline: recommending similar videos via randomly choosing a few videos in the same industry
- Audio features outperforms naïve baseline 94% of the time
- Audio features is be implemented in the first version of Peggy Olsen but GroupM is keen to incorporate it in the near future.

Discarded Features

Transcript

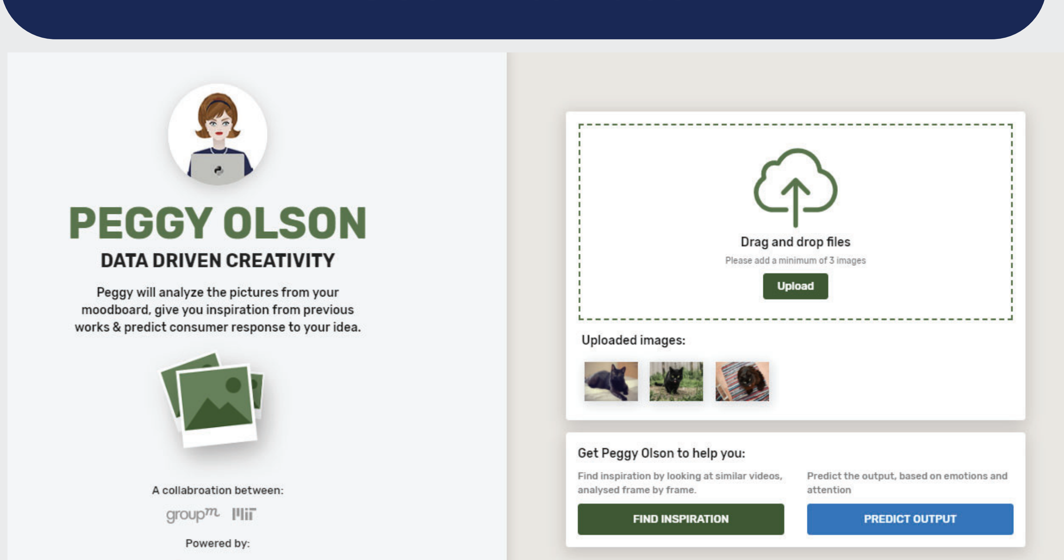
We implement Bag of Words and Word2Vec to embed the transcript. Transcript features are dropped because of the data accuracy. However, we expect that transcript will provide meaningful input if Google API can be more accurate in the future.

Web Entity & Logo

These two features are discarded due to inaccurate detection.

UI, Business Impacts & Next Steps

User Interface



Business Impact

1. Closed beta test: We ask 6 stakeholders within GroupM to upload 3-5 sets of mood board pictures into the tool and get back to us with whether they find the output recommendations are relevant. All test participants get inspirations from the recommendations and they find 80% of the videos are very relevant to their input pictures.
2. Time saved: GroupM estimates this tool can reduce the searching time that creative time spend on searching relevant videos for 159 hours in just one office after put into use.
3. Budget saved: GroupM expects to save 189,000 DKK in budget in Denmark office, with the feasibility to scale to the whole Nordic region.

Next Steps

1. Add More Videos into Database: expanding the video dataset allows Peggy to provide more relevant and inspiring results.
2. Improve the Tool after Gathering Feedback from Users: the bulit-in button will gather from the creative team once they put the tool into massive internal usage in 2020.
3. Improve Feature Accuracy: In the future, GroupM will increase feature accuracy by either working with Google and utilizing their improved API to re-extract features or developing their own tool.
4. Include Video Uploading Feature: GroupM will also include the option for the creative team to upload any sample video that they want to find similar videos with and in that case, the audio features will also be included for the recommendation criteria.