Case Study 1

By: Tobias Chin, Ryan Dalton, Jacob Ruppert, Rohan Prasad, and Bishoy Hanna

**1 Introduction:**

Using the News API we collected data on the titles of 219 ESPN articles along with the titles and headlines associated with it. With these findings, we hope to use image processing and other forms of analysis to better understand what sports and trends are most commonly represented in ESPN articles and using those findings we wish to make recommendations to advertisers.

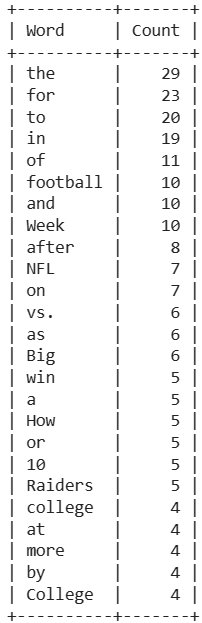
**2 Motivation:**

Our group decided to dive into football articles because it is something we were all passionate about. Football articles especially from Espn are read by millions of readers and the way companies intrigue us to read on. The titles of the articles act as a hook for readers, titles reference the main idea of the article in a certain way to grab the reader’s attention. Additonally, images associated with these articles provide the reader with a sense of familiarity and excitement. So we wanted to examine the tactics companies use to keep the reader interested through patterns using image processing and analysis.

**3 Analyzing the Data:**

**Data:**

**Most frequently used words:**



We pulled all individual words from the titles of the 219 articles using a for-loop.

Then used the PrettyTable package to produce a table of the most frequently used words.

**Model: Hugging Face Facebook/detr-resnet-50**

From source:

“The DETR model is an encoder-decoder transformer with a convolutional backbone. Two heads are added on top of the decoder outputs in order to perform object detection: a linear layer for the class labels and a MLP (multi-layer perceptron) for the bounding boxes. The model uses so-called object queries to detect objects in an image. Each object query looks for a particular object in the image. For COCO, the number of object queries is set to 100.

The model is trained using a "bipartite matching loss": one compares the predicted classes + bounding boxes of each of the N = 100 object queries to the ground truth annotations, padded up to the same length N (so if an image only contains 4 objects, 96 annotations will just have a "no object" as class and "no bounding box" as bounding box). The Hungarian matching algorithm is used to create an optimal one-to-one mapping between each of the N queries and each of the N annotations. Next, standard cross-entropy (for the classes) and a linear combination of the L1 and generalized IoU loss (for the bounding boxes) are used to optimize the parameters of the model. “ -@article{DBLP:journals/corr/abs-2106-00666,

author = {Yuxin Fang and

Bencheng Liao and

Xinggang Wang and

Jiemin Fang and

Jiyang Qi and

Rui Wu and

Jianwei Niu and

Wenyu Liu}

**4 Findings:**

When viewing the most frequently used words from article titles we can identify sport specific words, terms relating to American Football are some of the most frequently used:

| Word | # used | What does it mean? |
| --- | --- | --- |
| football | 10 | As ESPN is a US based company this term likely relates to American Football, but it may also refer to soccer. |
| Week | 10 | As ESPN focuses on sports articles, 'Week’ likely refers to specific weeks of sports seasons; this terminology is used frequently for weeks of the NFL season. |
| NFL | 7 | This term explicitly relates to American Football. |
| vs. | 6 | In the context of sports this term relates to one-on-one or team-based sports. |
| Big | 6 | The capitalized word ‘Big’ is most often associated with the Big 10 of American College Football in the context of sports. |
| win | 5 | This word's relevance to sports is self explanatory. |
| 10 | 5 | Along with the word ‘Big’ this number is most often associated with the Big 10 of American College Football in the context of sports. |
| Raiders | 5 | This is a NFL team. |
| college | 4 | While college could refer to any number of sports, American College Football is consistently one of the most popular collegiate sports programs. |
| College | 4 | As previously stated, while College could refer to any number of sports, American College Football is consistently one of the most popular collegiate sports programs. |

**5 Analysis:**

**Of the images from the 10 most popular articles…**

* 9/10 are American Football
  + 5 of those are NFL and 4 are College Football
* 4/10 prominently feature Quarterbacks
* 4/10 feature a football

****

* 2/10 feature Michigan State

****

* 1 featured the Raiders, who were also the most mentioned NFL team based on words used in titles

****

**Key Takeaways:**

* Star player or someone well known on the title, usually with football in hand
* Big 10 College Football and the NFL are relatively comparable in their popularity.
* Quarterbacks are the most popular position.
* Footballs are often included in the images

**Concerns:**

Why is Michigan overrepresented?

* Michigan Football is currently awaiting a potentially large fine from the NCAA.

Why are the Raiders mentioned so much?

* The Raiders just fired their controversial head coach.

-While these teams make headlines, they might not be advertiser-friendly.

**6 Business questions and ideas:**

**Business Question:**

How can a company leverage NewsAPI data from ESPN articles to identify emerging sports trends and capitalize on advertising opportunities in the evolving landscape of sports content consumption?

**Description of the Business Case:**

The business case involves using NewsAPI data, specifically from ESPN, to proactively identify emerging sports trends and leverage this information for strategic advertising planning. The aim is to stay ahead of the curve, understand shifting consumer interests, and capitalize on advertising opportunities in the dynamic landscape of sports content.

**How NewsAPI Data Could Help Allocate Resources:**

**Trend Identification:**

Analyze article titles and headlines to identify emerging sports trends and topics gaining popularity.

Use this information to anticipate shifts in consumer interests and allocate resources toward advertising in sports areas with growing engagement.

**Content Relevance:**

Evaluate the relevance of current advertising content in relation to emerging sports trends.

Tailor advertising messages to align with the latest trends, ensuring that the content remains fresh and resonates with the evolving interests of the audience.

**Targeted Campaigns:**

Identify niche or up-and-coming sports, teams, or players gaining traction in ESPN articles.

Develop targeted advertising campaigns to capitalize on the rising popularity of these specific sports or entities.

Allocate resources strategically to maximize the impact of campaigns in growing segments of the sports market.

**Platform and Format Optimization:**

Determine the preferred platforms and content formats for consuming emerging sports content.

Allocate resources to optimize advertising for the platforms and formats most popular among the audience engaging with new sports trends.

**Collaborative Partnerships:**

Identify potential collaboration opportunities with emerging sports stars, teams, or leagues.

Allocate resources towards partnerships that align with growing trends, enhancing the brand's association with up-and-coming sports content**.**

**Agility in Campaign Planning:**

Develop an agile advertising strategy that can quickly adapt to changes in sports content consumption trends.

Use NewsAPI data regularly to stay informed about evolving patterns, enabling the company to adjust advertising plans in real-time.

**7 Conclusions:**

**Based on our findings…**

We recommend that advertisers who hope to appeal to the ESPN demographic focus on Football in their advertisements. Football was overwhelmingly the most represented sport in the data set as a notable amount of the title words related to it and 9/10 top articles were football related.

Quarterback's are especially popular and most images with one included them holding a football, so an advertisement with both of those would be optimal. Whether they’re collegiate or professional doesn’t matter as long as the college players are from the Big 10 as it was mentioned repeatedly in the titles and Big 10 teams were the only college teams shown in the images.

Finally, avoid Michigan and the Raiders as both are currently involved in controversies. Controversy is good for headlines but isn’t as useful for advertisers.

**8 Citations:**

“News API – Search News and Blog Articles on the Web.” *News API: Search News and Blog Articles on the Web*, [News API – Search News and Blog Articles on the Web](https://newsapi.org/). Accessed 1 Nov. 2023.