TEXT ANALYSIS AND NATURAL LANGUAGE PROCESSING

Proposal idea for the final project – Leona Hasani

Working Title

Guiding Radiance: Navigating Ingredients and User Experiences in SPF Face Creams for Enlightened Skincare Choices (the working title will evolve within the evolvement of the project).

Basic Research Question

To what extent do the advertised benefits of popular SPF face creams align with their actual formulations and real-world user experiences, and how can this information guide consumers in making informed choices about sun protection skincare?

Key Papers

The papers that I will be focusing are:

1. U. Osterwalder and B. Herzog. Sun protection factors: worldwide confusion, *British Journal of Dermatology*, Volume 161, Issue s3, 1 November 2009, Pages 13–24, https://doi.org/10.1111/j.1365-2133.2009.09506.x
2. Schalka S, Reis VMS. Sun protection factor: meaning and controversies. An Bras Dermatol. 2011;86(3):507-15
3. N. P. Arthamevia, Adiwijaya and M. D. Purbolaksono, "Aspect-Based Sentiment Analysis in Beauty Product Reviews Using TF-IDF and SVM Algorithm," *2021 9th International Conference on Information and Communication Technology (ICoICT)*, Yogyakarta, Indonesia, 2021, pp. 197-201, doi: 10.1109/ICoICT52021.2021.9527489.

In this paper, I will also use material from websites such as the New York Times newspaper and other popular beauty magazines. Then, I'll be web scraping the necessary data from the SkinCarisma website for the ingredients of each spf face cream, as well as from the official spf face creams websites, to extract the reviews and overall 'grade' for those particular items.

Motivation

As someone who's always been a bit skeptical about beauty trends, I'm diving into the world of SPF face creams for my final project. You see these 'magic' creams everywhere on social media,

in newspapers, and magazines. But do they really deliver on their promises? I want to dig deep, not just for myself but for everyone else wondering. As I go through the world of SPF face creams, my focus extends to the real experiences through sentiment analysis of user reviews. I aim to decode not just what's in the products but also how users genuinely feel about them, providing a nuanced perspective for smarter skincare choices.

Idea

Framing it as a hypothesis:

The sentiment analysis of user reviews for SPF face creams will expose a divergence between marketing assurances and genuine user experiences, suggesting that the proclaimed effectiveness may not consistently match the diverse realities encountered by users.

Data

In this research paper, I plan to acquire data from reputable sources such as the New York Times and beauty magazines to identify the most recommended SPF face creams of the year. Additionally, I will delve into the Skincarisma website to extract detailed ingredient information for each of these selected creams, aiming to discern commonalities in ingredients across a spectrum of products. Simultaneously, I will collect user reviews directly from the official websites of these products to gain qualitative insights into people's experiences. To ensure a well-rounded understanding, I aim for a balanced sampling of both high-end and budget-friendly products. Potential obstacles, such as web scraping restrictions, may lead to a focus on products where data extraction is more feasible.

The goal is to integrate ingredient analysis with user sentiments, creating a comprehensive narrative that not only reveals ingredient trends but also provides real-world insights into the effectiveness and satisfaction levels of SPF face creams. Through this data-driven approach, the research seeks to empower readers with practical guidance on selecting SPF products effectively, addressing the overarching question of whether the hype surrounding these products aligns with user experiences and if price correlates with ingredient quality.

Tools

Based on the skills developed in this course, I intend to use Python as the primary tool for data extraction, cleaning, and analysis in the SPF face cream project. Web scraping methodologies, such as Beautiful Soup and Scrapy, will be used to extract insights from official product websites and Skincarisma. I intend to use my knowledge of NLP from the course to perform advanced sentiment analysis utilizing python libraries for decoding of user evaluations. This project provides a significant opportunity for practical application and growth of learned competence,

with a continuing commitment to smoothly incorporate evolving knowledge throughout the semester.

What's New

Unlike other studies just looking at what's in SPF face creams, this paper will take another fresh angle. I'll not just list the ingredients; I'll also dig into what people are saying about these creams. It's like getting the full story – not just the facts, but also how real people feel about using these products in their daily lives.

So, What?

Why does this matter?

Well, it helps bridge the gap between what companies promise and what actually happens. I'll not just tell you what's inside; I'll also share how people really feel about these popular face creams and how do they rate them. This kind of information is gold for anyone wanting to make smart choices in their skincare routine.

Contribution

The big contribution is giving people the whole picture. I'll not only talk/write about ingredients but also bring in real people's experiences. This way, you can make choices based on both the facts and how these products actually work for you. It's about making spf facial skincare decisions less confusing and more straightforward for everyone.