VEGGIE MINGLE

ADVANCED DATA MINING TO IMPROVE MATCHING EFFICACY

VEGGIEMINGLE: NAVIGATING THE LANDSCAPE OF PLANT-BASED LOVE

In the bustling world of online dating, niche platforms have always found a way to stand out. One such platform, VeggieMingle (a fictitious platform), emerged in the late 2010s, catering to the growing community of vegetarians, vegans, and other plant-based diet enthusiasts in the US. As the trend towards plant-based diets grew, so did the desire for like-minded partners who shared similar ethical, environmental, and health perspectives.

By 2023, VeggieMingle had established itself as the go-to platform for vegetarians and vegans seeking companionship. However, the journey wasn't always smooth. The dating industry was fiercely competitive, with giants like Tinder and Bumble dominating the market. Yet, VeggieMingle's unique value proposition allowed it to carve out a dedicated user base.

VeggieMingle faces a challenge: Despite a dedicated user base, match success rates aren't meeting expectations. Harnessing user data more effectively could be the key to fostering more meaningful connections on the platform.

In the spacious conference room of VeggieMingle's San Francisco headquarters, two executives, Maya Johnson, the Chief Marketing Officer, and Rajiv Mehta, the Head of Data Analytics, sat down for a strategic discussion. The ambiance was tense, with the weight of the company's future palpable in the air.

"We've got a strong user base, but our match success rate isn't where we want it to be," Maya began, her eyes scanning the data charts spread out in front of her. "We need to ensure that our users not only stay on our platform but also find meaningful connections."

Rajiv nodded, "I've been looking at the data, and while we have a plethora of information on our users, we're not leveraging it effectively. We must understand our users better dive deeper into their preferences, habits, and behaviors."

Maya leaned forward, "That's precisely what I was thinking. We need to turn this data into actionable insights. If we can predict what factors lead to successful matches, we can better guide our users, maybe even tweak our algorithms."

Rajiv added, "We've got data on diet types, reasons for their dietary choices, activity preferences, and so much more. There's a goldmine of information here. We need to dig in."

The conversation shifted to the challenges they faced. Maya highlighted the feedback they'd received. "Users often mention they want more than just a shared diet preference. They're looking for shared values, hobbies, and lifestyles. We need to consider all these factors."

Rajiv, ever the data enthusiast, was already thinking ahead. "We could potentially set up a classification problem. Using various advanced data mining methods, we could predict the likelihood of a match based on various factors. This could give us insights into what's working and what's not."

Maya smiled, "That's the spirit! Let's get to work. VeggieMingle has always been about connecting people. It's time we take it to the next level."

Approach and Expectations

As you delve into the VeggieMingle case, approach it as a data consultant tasked with revolutionizing the user experience on the platform. With the rich dataset, harness the advanced data mining techniques you've learned to uncover hidden patterns, relationships, and insights. Your goal is not just to analyze but to translate these insights into actionable strategies that can enhance the matching algorithm and overall user experience. Remember, while the data provides a roadmap, it's the nuanced understanding of user behaviors and preferences that will drive success. As you conclude, provide clear, data-backed recommendations that VeggieMingle can implement. Think holistically, considering both the immediate impact and long-term sustainability of your proposed solutions.

Your submission will be a one-page write-up (plus a cover sheet). Students may use appendices as necessary to justify their proposed solutions (and I strongly encourage you to use such appendices so that I believe your analysis). Append your R-code to the end of the submission. This is an individual assignment.

Data Dictionary

Variable Name	Description	Type	Possible Values
User ID	Unique identifier for each user	Numeric	1, 2, 3, ...
Age	Age of the user	Numeric	18-50
Gender	Gender of the user	Categorical	Male, Female, Other
Location	City where the user is located	Categorical	New York, Los Angeles, Chicago, ...
Diet Type	Type of vegetarian diet the user follows	Categorical	Vegetarian, Vegan, Raw Vegan, Pescatarian
Diet Duration (years)	Number of years the user has been following their diet	Numeric	1 to 20
Reason for Diet	Main reason the user follows their diet	Categorical	Ethical, Health, Environmental, Religious
Dietary Transition	How the user transitioned to their current diet	Categorical	Gradual, Overnight, In Phases
Favorite Cuisine	User's favorite type of cuisine	Categorical	Indian, Mediterranean, Asian, Western
Cooking Skill	User's cooking skill level on a scale of 1-10	Numeric	1 to 10
Exercise Habits	How often the user exercises	Categorical	Daily, Weekly, Rarely
Favorite Activity	User's favorite leisure activity	Categorical	Reading, Hiking, Yoga, Music
Alcohol Consumption	User's alcohol consumption habits	Categorical	Regularly, Socially, Never
Smoking Habits	User's smoking habits	Categorical	Regularly, Socially, Never

Education Level	Highest level of education achieved by the user	Categorical	High School, Bachelor's, Master's, PhD
Occupation	User's profession	Categorical	Engineer, Doctor, Teacher, Artist, Entrepreneur
Income Level	User's income bracket	Categorical	Low, Middle, High
Relationship History	Number of past relationships the user has had	Numeric	0-5
Children	Whether the user has children	Categorical	Yes, No
Pets	Whether the user has pets	Categorical	Yes, No
Travel Preference	User's travel style preference	Categorical	Adventurous, Relaxing, Cultural, Rarely Travel
Music Taste	User's favorite type of music	Categorical	Rock, Jazz, Classical, Pop
Movie Preference	User's favorite type of movie	Categorical	Action, Romance, Documentary, Comedy
Reading Habits	Number of books the user reads in a year	Numeric	1 to 20
Online Activity	Hours the user spends online daily	Numeric	1 to 10
App Usage Frequency	How often the user uses the VeggieMingle app	Categorical	Daily, Weekly, Monthly
Event Attendance	How often the user attends VeggieMingle events	Categorical	Often, Rarely, Never
Referral Source	How the user heard about VeggieMingle	Categorical	Friends, Online Ad, Social Media
Account Duration	Number of months the user has been on VeggieMingle	Numeric	1 to 24
Profile Completeness	Percentage of the user's profile that is complete	Numeric	50-100
Response Rate	How often the user responds to messages on VeggieMingle	Categorical	Always, Often, Rarely
Safety Concerns	Whether the user has safety concerns using VeggieMingle	Categorical	Yes, No
Matched	Whether the user found a match within 30 days of joining	Categorical	Yes, No