The Case of Mislabelled Recipes: Building an Asian Recipe Classifier Using SVM and Multinomial Logistic Regression with TD IDF Vectorization

Fig. 1. Assorted Asian dishes. Image from "Free Asian Food Photos," Pexels, https://www.pexels.com/search/asian%20food/ (accessed Nov. 25, 2025).

Background: It's a cold, winter Friday night. It's been a long week and the thought of using your last ounce of energy to peel onions and thaw the chicken is weighing heavily on your chest. A light bulb ignites in your mind. It's a revolutionary idea you are having, but you dare not entertain it right away. It'll be bad for your bank account, the vegetables in your fridge are going bad, you just spent money and time grocery shopping on Wednesday. Another thought arises. *One night can't hurt. I deserve it.* It's game over. You grab the phone and dial. You know the number by heart. "Hello, can I place an order for take out. One order of General Tso's. Yes, thank you." You drive up to the sketchy part of town, lit with neon signs and an eerily empty parking lot. You come inside and grab your steaming hot, brown paper bag from the counter as a golden Buddah waves from the corner of the empty restaurant with tables that look untouched. In all of your time getting takeout, you have never seen anyone at the tables. Your phone rings. "Hey, yeah let's watch a movie. I just grabbed Chinese for dinner. See ya," you say. The truth is, you did not get Chinese food. You picked up an American invention and adaptation of Chinese cuisine that has no tie to China at all. Congratulations, you are having American for dinner. Don't fret, it's delicious and others think so too! In fact, 73% of 3,144 U.S. counties have at least one Asian restaurant [1]. Japanese, Chinese, Korean, and Thai foods are extremely popular in the United States. Additionally, 12% of American restaurants have Asian food on their menus [2]. With East Asian and Southern Asian cuisine being consumed nationally, it's important that we stay cognizant of which recipes are authentic and what region they originate from. Gyoza, for example, is widely available at Japanese restaurants around the United States, but this is originally a Chinese invention.

Task and Deliverable: You're a foodie data scientist tasked to investigate whether Asian recipes on American websites are being labeled correctly. You must build an Asian recipe classifier that gives credit where credit is due and successfully classifies recipes across Asian countries based on their ingredients. Your model will be successful when you achieve 70% accuracy on this mission! Everything you need is here: https://github.com/gatesz33/CS3-DS4002?tab=readme-ov-file.

[1] J. Scipioni, "Analysis: Asian food winning over America," The Food Institute, Jun. 9, 2023. [Online]. Available: https://foodinstitute.com/focus/analysis-asian-food-winning-over-america/. [Accessed: Sep. 10, 2025].

[2] P. Taylor, "71% of Asian restaurants in the U.S. serve Chinese, Japanese or Thai food," Pew Research Center, May 23, 2023. [Online]. Available: https://www.pewresearch.org/short-reads/2023/05/23/71-of-asian-restaurants-in-the-u-s-serve-chinese-japanese-or-thai-food/. [Accessed: Sep. 10, 2025].