Food Quality and Preference 98 (2022) 104530

Contents lists available at ScienceDirect

Food Quality and Preference

journal homepage: www.elsevier.com/locate/foodqual

Exploring public perceptions on alternative meat in China from social media data using transfer learning method

Yuan Chen *, Zhisheng Zhang *

School of Mechanical Engineering, Southeast University, Nanjing 211189, China

ARTICLE INFO

Keywords:
Alternative meat
Consumer attitudes
Sentiment analysis
Machine learning
Data science
Topic model

ABSTRACT

The emerging social media serves as a complementary source for consumer behavior analysis with spontaneous data it generates. However, most studies employ time-consuming content analysis or lexical sentiment analysis. Considering the richness of data and progress of data science, in this paper, we propose a transfer learning based method to explore public attitudes towards alternative meat (AM) using data from social media in China to provide an alternative perspective. We compare traditional machine learning models: Naive Bayes and Support Vector Machine with our BERT-based Alternative Meat (BAM) model on the annotated sample. BAM model outperforms others in terms of macro F1 score and accuracy and is employed on the whole dataset later. The sentiment analysis result shows that among 41782 related posts we accumulated, about 42.10% of posts are personal posts consisting of negative, neutral, and positive feelings towards AM with a proportion of 28.77%, 22.91%, and 48.32% respectively. It is less promising compared with the consensus previous studies reach that over half of the Chinese people are positive and few Chinese are negative towards AM. Our findings add to the blooming body of studies suggesting the relationship of people's willingness to try or purchase AM and factors including gender, geography, price, veganism, and food safety. Conspiracy theory is identified for the first time as the main reason for opposition to AM among Chinese consumers. Instead of the booster, traditional vegetarian substitutes especially tofu turn out to be an obstacle for accepting AM with much resemblances.

1. Introduction

1.1. Literature review of alternative meat

People have been eating meat from animals: cows, chickens, and pigs for so many years and farm-grown meat has been the human's main source of protein for generations. Due to environmental problems (Aleksandrowicz, Green, Joy, Smith, & Haines, 2016) and health awareness (Tilman & Clark, 2014), people have turned their eyes to alternative meat which is more eco-friendly and healthy. The alternative meat could be mainly divided into two categories: plant-based meat and cultured meat. Plant-based meat is made from vegetable proteins like soy, peas, or wheat protein with amino acid, fat, and other ingredients to add flavor. Cultured meat is made by growing totipotent stem cells or muscle cells in the cultivator through "tissue-engineering" technology. International companies are heavyweights in the new wave of promising novel meat-alternative while many startups are ready to steal a march on the bigger companies. However, despite the rush and enthusiasm of scientists, enterprises and media, it is the customers' attitude that will decide the future of alternative meat.

There are already many studies on public acceptance of alternative proteins like plant-based meat and cultured meat but mainly in western countries (Bryant & Barnett, 2020; Hartmann & Siegrist, 2017; Onwezen, Bouwman, Reinders, & Dagevos, 2021). Most studies focus on willingness to pay (WTP) and the reason behind it. Gómez-Luciano, de Aguiar, Vriesekoop, & Urbano (2019) find that customers prefer plant-based meat to other alternative proteins and there is a difference in attitudes among dissimilar economical countries. Siegrist and Hartmann (2020) discover that there is a substantial difference in acceptance of cultured meat across countries and culture plays an important role in people's decision-making towards cultured meat. The viewpoint that willingness differs across countries is in line with findings from other studies (Chriki & Hocquette, 2020; Grasso, Hung, Olthof, Verbeke, & Brouwer, 2019). Bryant and Sanctorum (2021) find that product quality like taste is the people's top concern for alternative meat. They also find that people who are female, young and vegetarians tend to be more satisfied with existing alternative meat. Slade (2018) find that people prefer beef burgers to plant-based and cultured meat burgers given the

* Corresponding authors.
E-mail addresses: 230159322@seu.edu.cn (Y. Chen), oldbc@seu.edu.cn (Z. Zhang).

https://doi.org/10.1016/j.foodqual.2022.104530
Received 8 August 2021; Received in revised form 25 December 2021; Accepted 5 January 2022
Available online 10 January 2022

same price based on a hypothetical choice experiment. Besides price sensitivity, they also find that people who are vegetarians, younger, and more educated are more likely to choose simulated meat rather than beef ones while opinions of other food technologies, environment and agriculture all have an influence on choice making too. Wilks, Phillips, Fielding, and Hornsey (2019) explore the psychological reason behind people's attitudes towards cultured meat and find that food neophobia is the strongest predictor. According to literature results, flavor, prices, and safety seem to be the public's top concerns. Gender, veganism, age, income, and cultural difference all play an important role in people's choice of alternative meat.

China has a long history of vegetarian substitutes i.e. soya products for centuries long. Different varieties of traditional mock meat have been a daily part of the culinary scene in China. This food culture brings much uncertainty to how Chinese will react to modern alternative meat. Considering the different cooking styles, economic level, and culture from western countries, China, as a huge potential market for global meat consumption with increasing meat demand (Godfray et al., 2018), is worth a deep research on the latest public attitudes towards alternative meat.

So far, only limited academic research works have been undertaken on Chinese acceptance of alternative meat. Dempsey and Bryant (2020) explore how different descriptions of cultured meat affected levels of acceptance and the most favorable product format of cultured meat in China. Bryant, Szejda, Parekh, Desphande, and Tse (2019) find a similar pattern for clean meat and plant-based meat purchase likelihood within China, India, and the USA although the willingness from China and India are sustainably higher than that in the USA. Siegrist and Hartmann (2020) mainly focus on comparison among ten countries including China in terms of acceptance of cultured meat. Zhang, Li, and Bai (2020) examine consumer awareness and acceptance of artificial meat then explore the change in attitudes after obtaining basic information on artificial meat by questionnaire surveys. Liu, Hocquette, Ellies-Oury, Chriki, and Hocquette (2021) explore the Chinese attitudes towards cultured meat and find that the main concerns are about safety and unnaturalness rather than ethical and environmental issues. All the papers are based on surveys and tackle public acceptance of cultured meat particularly except Bryant et al. (2019). Apart from above survey-based papers, no further related research has been found.

1.2. Emerging social media data

All previous studies use either survey or historical documents data. They mainly collect data from participants by conducting a survey then perform factor analysis methods to identify the correlation between the WTP for alternative meat and other possible factors like gender, age, or income for further research. But the survey data are often time and space limited (Dorce, da Silva, Mauad, de Faria Domingues, & Borges, 2021). The subjective judgments play a vital role in customer satisfaction but there have been subjectivity concerns about honesty and accuracy of responses because of the external pressure the respondents receive from media, peers, and society (Frank, Cebrian, Pickard, & Rahwan, 2017). The problem of respondent honesty and accuracy during survey often leads to the potential bias and errors in survey experiments.

The emerging social media generates big data reflecting users' activities and thoughts (Ghani, Hamid, Targio Hashem, & Ahmed, 2019), and could serve as a complementary data source to food-related information seeking, consumer behavior analysis and product development (Carr et al., 2015; Kuttschreuter et al., 2014). The social media data are potential not only for its richness, and usability but the automatic retrieval of user-generated content representing the technological advance for consumer behavior analysis (Martí, Serrano-Estrada, & Nolasco-Cirugeda, 2019). The diversity of social media data, and the content retrieved from them all contribute to a multi-perspective approach to the study of consumer analysis. The dynamic data complete traditional survey method by providing more insights by tracking emerging trends, temporal patterns, and spatial-temporal patterns continuously at different spatial and temporal scales especially for places where systematic survey is not available (Alaa, Alhuwail, Househ, Hamdi, & Shah, 2020; Guntuku et al., 2020; Heikinheimo et al., 2017). In the food field, the prospective of employing data science including nature language processing (NLP) for text data in food data analysis is promising (Hamilton & Lahne, 2020; Moranges, Rouby, Plantevit, & Bensafi, 2021). Text analysis of user-generated text from social media like Twitter, online reviews and forums can deliver us valuable insights about consumer attitudes and behaviors (Jaeger & Rasmussen, 2021). Food scientists could explore state-wide nutrition behavior through Twitter (Abbar, Mejova, & Weber, 2015), food communication on Reddit among different healthy scale cities (Blackburn, Yilmaz, & Boyd, 2018), COVID-19 lockdown influence on food priorities through Google trends (Laguna, Fiszman, Puerta, Chaya, & Tárrega, 2020), consumer beliefs regarding organic food through online comments (Danner & Menapace, 2020), "gluten-free" topic information sharing on Twitter (Puerta et al., 2020), and sentiments towards new food trends from Twitter data (Pindado & Barrena, 2020). Compared with survey methods, the social media data method could deliver us current customer trends with big data dynamically and get real-time feedback which is very vital in this fast-changing world. However, no research has addressed the social media data exploration of consumer attitudes towards alternative meat before.

Current food data studies on social media lag behind at data science level. Most studies focus on simple word counting method based on frequency and occurrence (Carr et al., 2015), content analysis (Danner & Menapace, 2020; Jaeger & Rasmussen, 2021; Vidal, Ares, & Jaeger, 2016) which is very time-consuming especially for large amount of data, or a combination of both (Vidal, Ares, Machín, & Jaeger, 2015). Other studies investigate the usage of semantic networks (Grebitus & Bruhn, 2008), co-occurrence network (Puerta et al., 2020), and concept mapping approach (Peschel, Kazemi, Liebichová, Sarraf, & Aschemann-Witzel, 2019) for the analysis of associations and communications. Also, there are novel points like considering the use of emoticon and emoji in evaluating emotions (Jaeger, Roigard, & Ares, 2018, Jaeger, Vidal, & Ares, 2021; Jaeger, Lee, et al., 2017; Jaeger, Vidal, Kam, & Ares, 2017; Jaeger & Ares, 2017; Vidal et al., 2016, Vidal, Ares, Blond, Jin, & Jaeger, 2020) for tweets and open-ended questions. Some papers do explorations of assessing public sentiment towards food-related issues through social media data, but all employ lexicon-based sentiment analysis method, dividing text data into negative, neutral, and positive ones coarsely (Pindado & Barrena, 2020; Samoggia, Riedel, & Ruggeri, 2020; Tian, Lu, & McIntosh, 2021). Literature using advanced method like machine learning or deep learning to analyze food-related consumer sentiment with social media data in a fine-grained manner is relatively scarce.

1.3. Research aim and empirical overview

The present work aims to explore advanced data mining techniques on a fine-grained customer attitudes analysis of alternative meat with the combination of massive social media data and transfer learning method. We perform sentiment analysis with state-of-the-art (SOTA) method and provide an alternative perspective for assessing public attitudes on alternative meat through social media data in China. The workflow is illustrated in Fig. 1. First, we accumulate online data from an online social media called Weibo through web crawling by python using a combination of keywords. After performing data preprocessing procedures, we compare several machine learning and transfer learning models on the annotated sample posts with evaluation metrics of accuracy and macro F1 score. The best-performed model is then selected to classify whole posts into two general types: non-personal and personal posts with six categories to get the fine-grained classification results for final discussion. We perform factor analysis to find potential correlations between the public's attitudes and factors like gender and areas. Text

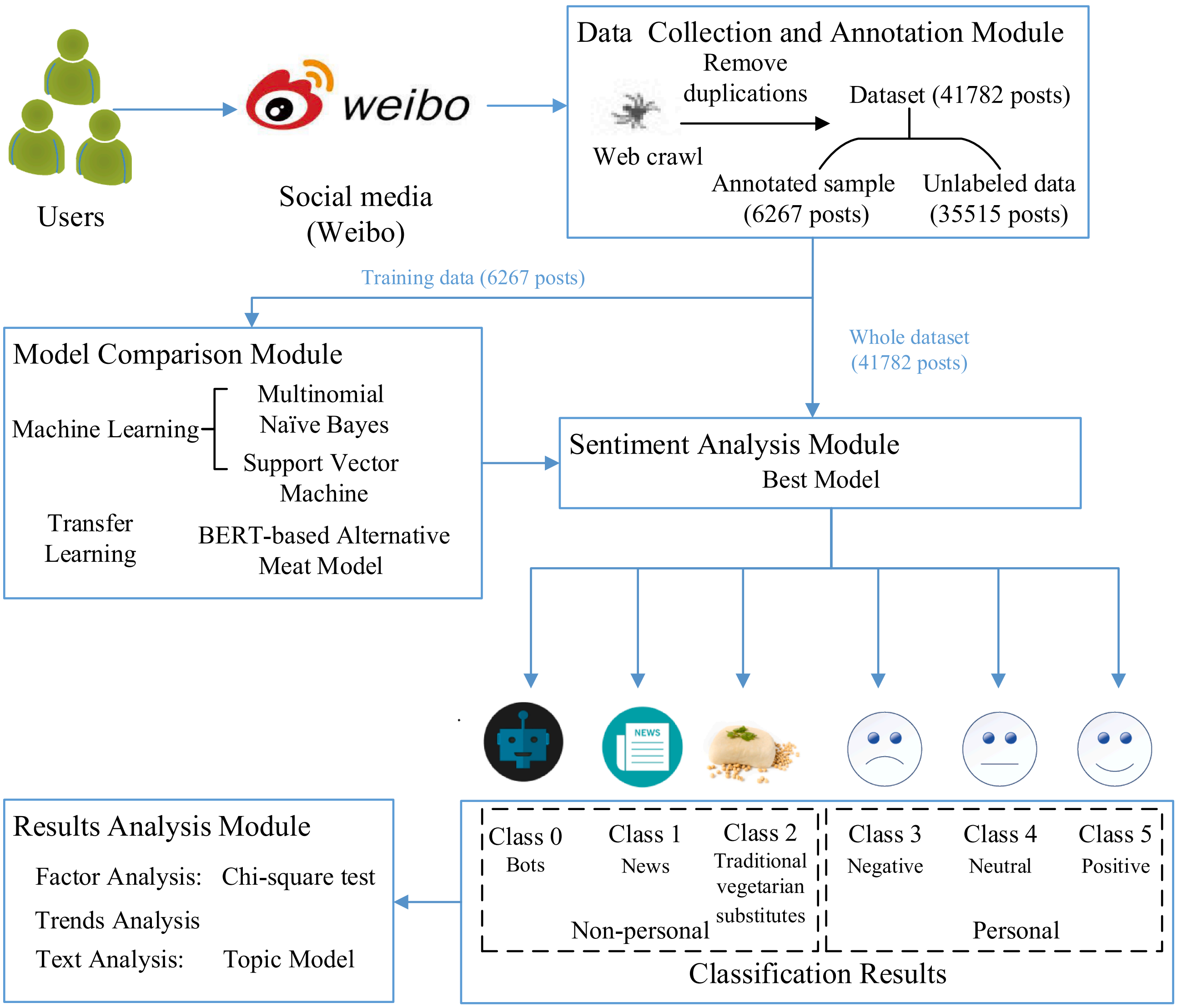

Fig. 1. The workflow.

analysis is utilized by Topic Model to extract topics to understand public interests and concerns for further analysis. This study could provide reference strategies for future data science research on customer attitudes analysis.

2. Material and methods

2.1. Data collection and sample annotation

We collect relevant data by web crawling using Python from Sina Weibo[1], the most active and Twitter-like social media platform in China. ("weibo" is a general term as the Chinese word for "microblog", but Sina Weibo is commonly called Weibo.) According to the Weibo's user development report for 2020[2], Weibo has about 511 million Monthly Active Users, 224 million Daily Active Users as a leading social media platform in China. People could share text, pictures, and videos on their own pages by sending out a weibo post. Posts (reposts included) from one's followings would appear in one's own timeline. The users could also repost a weibo post with the option of adding their own comments alongside, very similar to Twitter's Quate Tweet function. Although Weibo has an official API (Application Programming Interface), it does not provide "streaming" API tools commonly used to obtain random samples over time like Twitter. There is an advanced search functionality of Weibo that people could obtain weibo posts within certain topic, time, area and other settings. So we decide to use web crawler to scrape web information automatically instead. The crawler starts from one URL (Uniform Resource Locator), accesses all the URLs associated with it, and obtains and extracts the valuable data from each page. As the crawler visits these URLs: the search pages within certain search restrictions with Requests[3], the Internet returns the HTML (Hyper Text Mark-up Language) file to the downloader, which saves the file locally. We use lxml[4] as HTML parser to obtain posts' information like content, publish time eventually. We choose '植物肉'(*zhiwurou*), '植培肉'(*zhipeirou*), '人造肉'(*renzaorou*), meaning plant meat, plant cultivated meat, and man-made meat literally in Chinese as search words since these are

[1] https://www.weibo.com/

[2] https://data.weibo.com/report/reportDetail?id=456

[3] https://pypi.org/project/requests/, Requests is a HTTP (Hyper Text Transfer Protocol) library in python, could send HTTP requests to specified URLs and return the Response Object

[4] https://pypi.org/project/lxml/, lxml is a python library for processing XML (Extentsible Markup Language) and HTML(Hyper Text Mark-up Language)

the most commonly used names for plant-based meat and cultured meat in China. Most official organization and companies use these terms in news coverage and commercial promotions in China.

Once we finish the accumulation, duplicated posts are removed by their unique id crawled alongside. Finally, we have accumulated 41,782 related weibo posts (both original and repost ones are included) published from January 1, 2020, to April 30, 2021. The Weibo posts information we extract including id, content, publish time, publisher's name, location, and gender. A full description and trends illustration of the whole dataset are shown in Table 1 and Fig. 2. We select a sample of 6,267 posts from the whole dataset randomly. In Chinese, the word '人造肉' or '植物肉' could also be referred to the traditional Chinese mock meat, so the collected weibo posts include posts talking about traditional vegetarian substitutes. There are many bots posts containing irrelevant content towards alternative meat too. Also, these posts encompass many non-personal narrative data like news which are not generated by individual users but organizations or official accounts, unsuitable for consumer behavior analysis.

Previous studies often divide posts, reviews or tweets into positive, neutral, and positive categories roughly ignoring the existence of irrelevant and non-personal data. Considering the complexity of accumulated posts mentioned above, we decide to have a more fine-grained classification by dividing sample tweets into six categories: class 0: bots, class 1: news, class 2: traditional vegetarian substitutes, class 3: negative attitudes, class 4: neutral attitudes, and class 5: positive attitudes after thorough analysis. The posts from class 0, 1, 2 belong to non-personal ones while the class 3, 4, 5 consist of personal posts. The distribution and classification of the sample dataset are displayed in Table 2. This dataset will be used in the following part for model training.

2.2. Model comparisons

We select several state of art methods to explore the classification of related Weibo posts. For sentiment analysis, a vital part of NLP, the traditional computational linguistics method is mainly unsupervised and uses sentiment lexicon to calculate the level of the sentiment of words, sentences, and documents. With the development of computer science, artificial intelligence has gained great success in many research fields. Supervised methods like machine learning and deep learning have accelerated the pace and became a major part of sentiment analysis. In this paper, we proposed a BERT-based alternative meat model to utilize the classification of alternative meat. Competitive models like Support Vector Machine (SVM), Multinomial Naïve Bayes (MNB) have been chosen as the baseline model for comparison. All models used in this paper are supervised leaning methods which require labeled data for training.

2.2.1. Support Vector Machine (SVM)

Support Vector Machine (Cortes & Vapnik, 1995) is a supervised learning model often used in data analysis for classification and regression analysis. It constructs a hyperplane or a set of hyperplanes in a high- or infinite-dimensional space for category separation. Among all hyperplanes that classify data, the hyperplane that has the maximum margin between classes achieving the lowest generalization error represents the best one. SVM could perform a non-linear classification using kernel function to map inputs into high-dimensional feature spaces. In our experiment, we use Radial Basis Function (RBF) as our kernel function. The quadratic optimization problem is illustrated in (1). The RBF is shown in (2). For a training set $(x_i, y_i), i = 1, \cdots, l$ where $x_i \in R_n$ and $y \in \{1, -1\}^l$:

$$\min_{w,b,\xi}\frac{1}{2}W^TW + C\sum_{i=1}^{l}\xi_i subjectivetoy_i\left(W^T\phi(x_i) + b\right) \geq 1 - \xi_i\xi_i \geq 0 \tag{1}$$

$$K(x_i, x_j) = exp\left(-\gamma\|x_i - x_j\|^2\right), \gamma > 0 \tag{2}$$

Where $W^T(w, b) \in R^n \times R$ are the control parameters, ξ_i is the slack variable, ϕ is a non-linear map for mapping the inputs from original space to feature spaces. Two variables: C is a positive parameter, γ is the kernel parameter varying to obtain the optimal performance.

2.2.2. Multinomial Naïve Bayes (MNB)

Naïve Bayes is a statistical classification method based on the Bayes theorem that uses probability statistics knowledge for classification (Eyheramendy, Lewis, & Madigan, 2003). It is a supervised learning requiring a labeled training dataset for probability calculation with attribute conditional independent assumption. The Naïve Bayes algorithm is simple, common yet effective. We choose Multinomial Naïve Bayes classifier in our experiment. The posterior distribution of an example x belonging to class c_i is illustrated in (3).

$$p(y = c_i|x) = \frac{p(y = c_i)p(x|c_i)}{p(x)} \tag{3}$$

where $p(y = c_i)$ is the prior probability of class c_i , and $p(x|c_i)$ is the class-conditional distribution of x that appears in class c_i .

2.2.3. Bidirectional Encoder Representations from Transformers (BERT)

Large-scale pre-trained models (PTMs) like BERT (Bidirectional Encoder Representations from Transformers) (Devlin, Chang, Lee, & Toutanova, 2019) and Generative Pre-trained Transformer (GPT) have accomplished great success in the artificial intelligence (AI) field. Compared with other self-supervised Pre-trained word embedding methods like word2vec, Glove, and Elmo which are all feature transfer methods, BERT assumes that the target task shares or has prior distributions of hyper-parameters with the source task. BERT is based on the stacked multi-layer bidirectional Transformers. The Transformer is an encoder-decoder structure that applies self-attention mechanism, modelling correlations between all words of the input sequence in parallel. It can grasp knowledge and information ranging from shallow features, syntactic features, to semantic features from massive data which would benefit various downstream tasks especially for complex ones like sentiment analysis and text generation.

For machine learning models (SVM and MNB), unlike English whose words are naturally separated by space, the Chinese text should be segmented by splitting a sequence of characters into words first. After cleaning procedures, feature selection work is needed to extract features. Finally, the features would be feed into different classifiers for training and testing. In this paper, first, we use jieba[5] (a Python Chinese word segmentation module) to perform word segmentation. Then, stop words are removed. After that, the feature extraction part is accomplished by using Term Frequency-Inverse Document Frequency (TF-IDF) as input features. We use TfidfVectorizer imported from scikit-learn[6] (a Python module for machine learning) for the calculation of TF-IDF. The extracted input feature of TF-IDF would be fed into SVM and MNB classifier models for further classification. We use the multinomial Naive Bayes classifier with sklearn.naive_bayes.MultinomialNB module and support vector machine classifier with sklearn.svm.SVC module both imported from scikit-learn package to conduct the experiment. The mechanism of TF-IDF is explained below.

Table 1
Overview of the entire dataset.

Characteristics	Post type		Gender		Total
	Original	Repost	Male	Female	
Number	31378	10404	21754	20028	41782
Percentage	75.10%	24.90%	52.07%	47.93%	100%

[5] https://github.com/fxsjy/jieba
[6] Website: https://scikit-learn.org, https://github.com/scikit-learn/scikit-learn

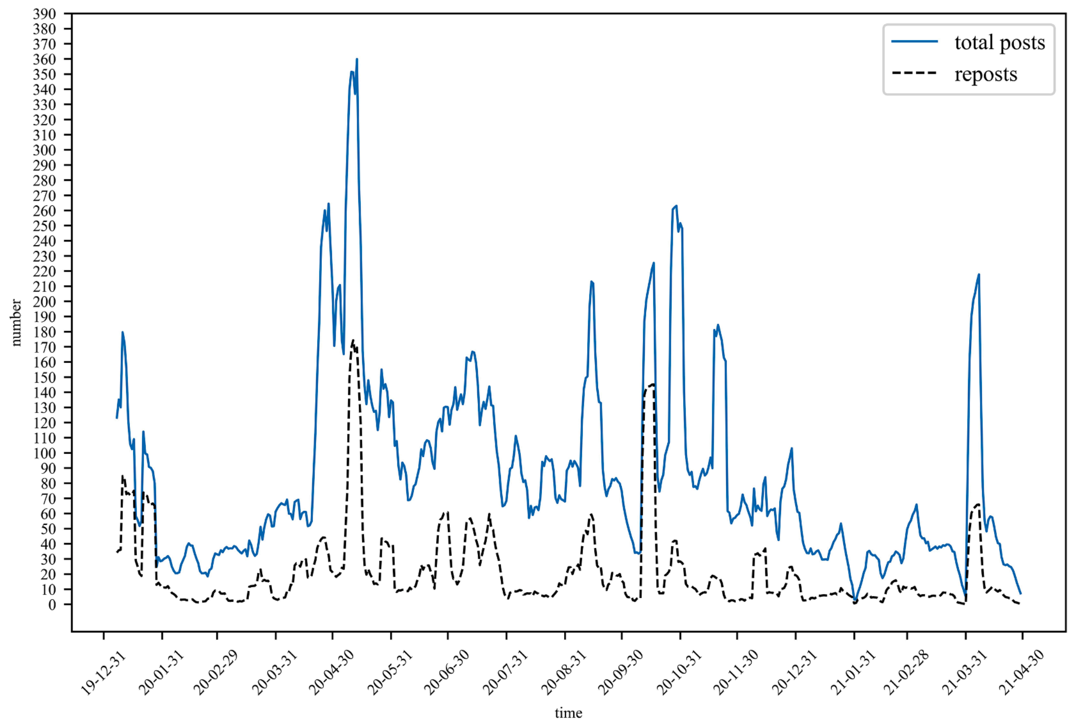

Fig. 2. The trends of related posts from January 1, 2020 to April 30, 2021. The blue line and dashed black line represent the rolling weekly average number of total posts and reposts during collected period respectively. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 2
Classification of the sample.

Classification type	Definition	Number
0: Bots	Nonsense, unrelated or commercial posts from bots	1122
1: News	Economic news and related reports	2133
2: Traditional Vegetarian Substitutes	About traditional plant-based alternatives like soy-products, tempeh and seitan	421
3: Negative	Express negative opinions like refusal, rejection, opposition or negative product evaluation	764
4: Neutral	Express unknown, unspecific, or neutral feelings	584
5: Positive	Express positive attitudes or willingness of trying alternative meat or satisfactory feedback on alternative meat products	1243
All	Total	6267

The TF-IDF is often used in text analysis and word scoring for natural language processing. It is calculated by multiplying two metrics to measure the importance of a certain term to the document in a set of documents. The score is proportional to term frequency which is the frequency certain word appears in a document. But the score decreases if the frequency of the word in the corpus increases which is the inverse document frequency means. For instance, in a smartphone related document, for words like 'battery' and 'the', the former one may have a lower term frequency in a document than the word 'the'. If only term frequency is used, the importance of the word 'the' would beat that of 'battery' which could go wrong. Hence, we introduce inverse document frequency to scale down the influence of these words which are common in all documents. The calculation of TF-IDF is shown in (4)-(6).

$$tf_{i,j} = \frac{n_{i,j}}{\left(\sum_k n_{k,j}\right)} \tag{4}$$

$$idf_{i,j} = \log\frac{D}{1 + D_{ti}} \tag{5}$$

$$tfidf_{i,j} = tf_{i,j} \times idf_{i,j} \tag{6}$$

Where $tf_{i,j}$ is the term frequency value of term i in document j, $n_{i,j}$ is the number of term i in document j. $idf_{i,j}$ is the inverse term document frequency of term i in document j, D is the document number in the corpus, D_{ti} is the number of document that contains term i.

For deep learning methods, we use state of art transfer learning method BERT as our pre-trained model for embedding layers. Such pre-trained learning models save much effort for feature extraction and could even skip above feature extraction part. In this paper, we fine-tuned BERT-Base, Chinese[7] (Chinese Simplified and Traditional version of BERT-Base) which contains 12 transformer block layers, 768 hidden units, 12 heads, 110 M parameters totally. The model consists of the fine-tuned BERT-Base, Chinese layer followed by a Batch Normalization (BN) layer (Ioffe & Szegedy, 2015), a Dropout layer (Srivastava, Hinton, Krizhevsky, Sutskever, & Salakhutdinov, 2014) to avoid over-fitting, and a Softmax layer for classifying posts into six different classes.

The architecture of BAM model is illustrated in Fig. 3. The BERT-Chinese takes sample size sequences of tokens with a maximum length of n as input. (To save computational cost, we choose the max length as 140.) Each word sequence can be represented as $w = \{w_1, w_2, \cdots, w_n\}$ where n denotes the length of sequence. The output of BERT, H consists of vector sequences with the length of 768 for each token in sentence. $\boldsymbol{H} = \{\boldsymbol{h}_1, \boldsymbol{h}_2, \cdots, \boldsymbol{h}_n\}$, where $\boldsymbol{h}_n$ corresponds to the nth token representation in each word sequence. So the shape of H is $\mathbb{R}^{n\times d}$, where $n = 140$, $d = 768$. We extract the output of the [CLS] token as the representation of sentence for sentence-level classification. Then the sequence will be fed into the BN layer, the drop out layer (dropout rate of 0.5) sequently, and finally the classifier of Softmax layer to get a sequence with the

[7] https://github.com/google-research/bert

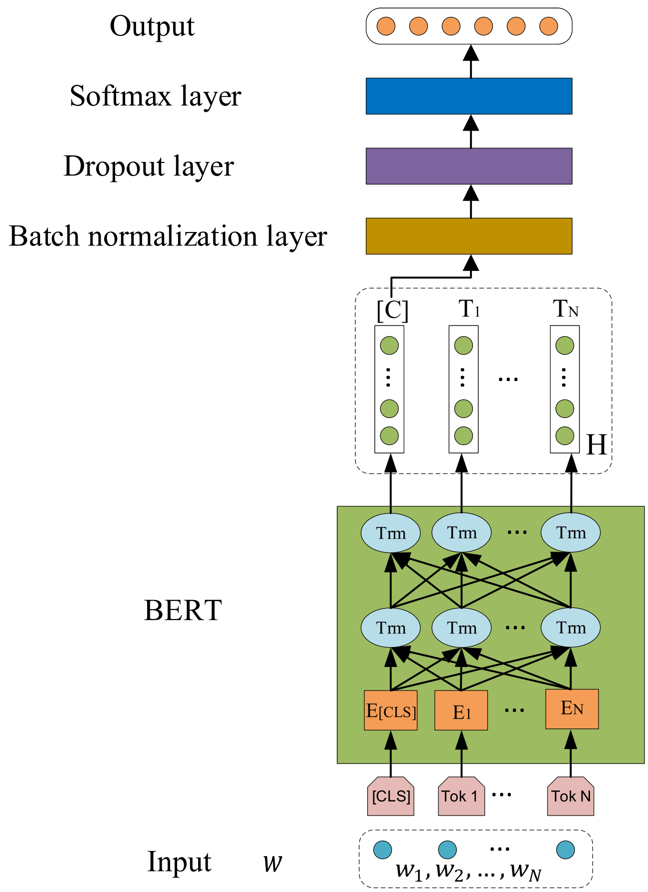

Fig. 3. The architecture of BERT-based Alternative Meat model. Trm represents transformer.

dimension of 6, the number of sentiment categories.

The content in the output array of Softmax layer represents the probability of multiple sentiment categories in the given text. The Softmax function is often used in the output layer for classification problems by predicting multinomial probability distribution. The probability of category can be obtained as shown in (7):

$$Softmax(c_i) = P(c_i|X,\theta) = \frac{P(X,\theta|c_i)\cdot P(c_i)}{\sum_{k=0}^{N} P(X,\theta|c_k)\cdot P(c_k)} \tag{7}$$

where $P(X,\theta|c_i)$ is the conditional probability of the sample X given class c_i , and θ is the model parameters. $P(c_i)$ denotes the prior probability of the class c_i .

Loss is calculated by the categorical cross-entropy loss function for multi-class classification as shown in (8). It is the cross entropy between estimated output probability distribution $Output_X$ and the target class probability distribution $Target_X$. The training goal of the model is to minimize the loss function and the parameter set is trained and updated through backpropagation algorithm. Other techniques to avoid overfitting like early stopping are also used in the experiments.

$$loss(Target_X, Output_X) = -\sum_{X} Target_X \log Output_X \tag{8}$$

where $Output_X$ is the estimated output probability distribution, $Target_X$ is the target class probability distribution, X represents the sample.

The Hyper-parameters obtained by grid search or fine-tuning are shown in Table 3. All codes are performed on Jupyter Notebooks on Colab (Tesla T4 GPU and 12G RAM) with python 3.7.10 environment. The SVM and MNB approaches are implemented within the Python scikit-learn package. The BAM model is implemented in python using Keras[8] with TensorFlow[9] as backend.

Table 3
Hyper-parameters setting.

Parameters	Value
Learning rate	1e-4
Dropout	0.5
Batch Size	32
Optimizer	Adam
C	128
Gamma	0.1

2.3. Evaluation

We perform 5-fold cross-validation to get the average results. We split the training data set into 5 folds, each fold serves as test corpus and the remaining four folds as training corpus in turn. For SVM and MNB methods, the classifier is trained on the training corpus and the average result of the 5 outcomes of the model evaluations on each test corpus is used as a predictor of the algorithm performance. For BAM model, we train the model on the training corpus, and test it on the test corpus after each epoch, this training process would stop if the evaluation metrics do not increase for 7 epochs. Then the model obtained the best scores on test corpus among all epochs would be considered as the best model. We conduct this process on each test fold and get the average results of all five best models.

Generally, scores like accuracy (ACC), Precision, Recall, and F1 score are used for the evaluation of text classification in natural language processing. F1 score is a combination of Precision and Recall score. Since it is an imbalanced multi-class scene in our case, that equal attention should be paid to each class, the macro F1 score is used. Macro F1 score gives each class equal weight for calculation regardless of its sample size, which avoids the miss of minority classes. So evaluation metrics consist of macro F1 score and accuracy score are used in this paper. The calculation of accuracy, macro F1 score is shown in (9)-(13).

For class c_i in a $(N+1)$ class set $\{c_0, c_1, \cdots, c_N\}$:

$$Precision_i = \frac{TP_i}{(TP_i + FP_i)} \tag{9}$$

$$Recall_i = \frac{TP_i}{(TP_i + FN_i)} \tag{10}$$

$$F1score_i = \frac{2Precision_i Recall_i}{Precision_i + Recall_i} \tag{11}$$

$$macroF1score = \frac{1}{N+1}\sum_{i=0}^{N} F1score_i \tag{12}$$

Table 4
Evaluation results.

Model	Accuracy	macro F1 score
Multinomial Naive Bayes	79.4639 % (±0.3392%)	76.2909 % (±0.9573%)
Support Vector Machine	85.7666 % (±0.5661%)	83.4908 % (±0.3302%)
BAM model	91.2558 % (±0.6987%)	89.4896 % (±0.7682%)

[8] https://github.com/keras-team/keras
[9] https://www.tensorflow.org

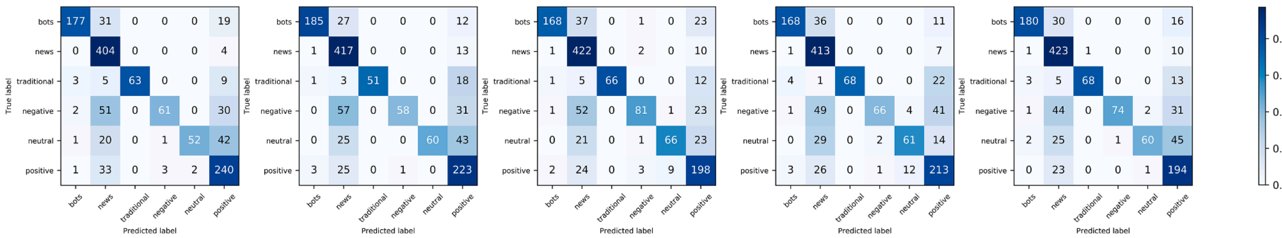

Fig. 4. The confusion matrix of MNB model.

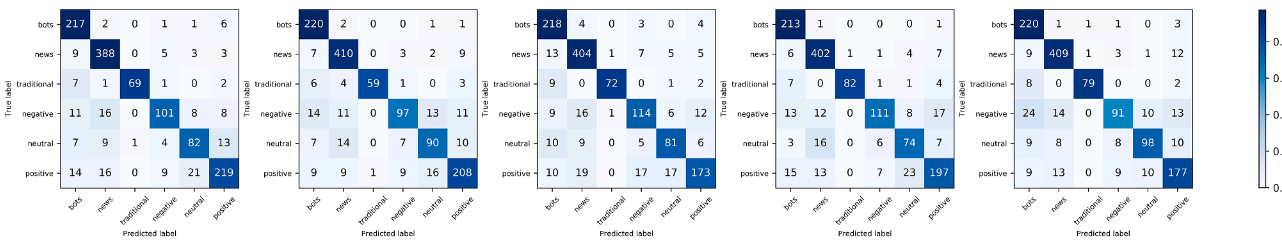

Fig. 5. The confusion matrix of SVM model.

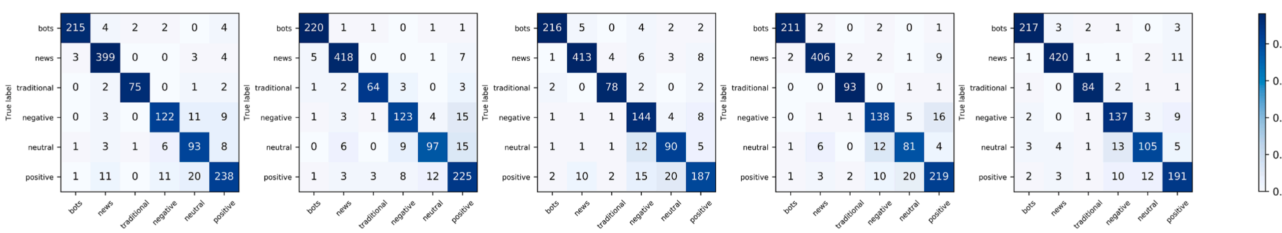

Fig. 6. The confusion matrix of BAM model.

$$ACC = \frac{\sum_{i=0}^{N} TP_i}{\sum_{i=0}^{N} (TP_i + FN_i)} \tag{13}$$

Where TP_i, TN_i, FP_i, FN_i are the true positive, true negative, false positive, false negative cases for class c_i , respectively.

The evaluation results are shown in Table 4. We can see that the transfer learning based model: BERT-based alternative meat model outperforms others regarding accuracy and macro F1 score. Also, the confusion matrix of five folds on test data of MNB, SVM and BAM models is shown in Fig. 4, Fig. 5, and Fig. 6 respectively. All three classifiers gain satisfactory results on classes of bots, news and traditional vegetarian substitutes but the performances start to diverge from personal posts significantly. Non-personal type posts have more distinguished linguistic characteristics compared with personal types which makes them easy to be picked out. The news often have different writing style compared with personal posts while bots, traditional vegetarian posts would differ much in content and relate less with other classes. Sentiment analysis is hard because the ambiguous way people talk and write. A same word could have different emotional orientations in different contexts. Online language processing is especially complex with the use of abbreviations, buzzwords, metaphor, and irony expressions. Compared with traditional models like SVM and MNB, BERT could grasp the deep features like hidden semantic features which could not be obtained through traditional feature extraction like TF-IDF. Besides that, the amount of class sample could also play a role in classification performance since more samples give the classifiers more semantic information to learn. Further studies could be data augmentations for biased and imbalanced dataset to improve the accuracy of minor classes.

2.4. Factor analysis

We conduct factor analysis to evaluate factors' relationship with public acceptance level based on sentiment results obtained from social media data. From previous analysis, we choose the publisher's gender and area as factors. To test the dependency between gender and sentiment category, and dependency between area and sentiment category, two null hypotheses are considered. Hypothesis 1: different gender group has the same frequency of sentiment category. Hypothesis 2: different geographic group has the same frequency of sentiment category. Hypothesis 1: different gender group has the same frequency of sentiment category. Hypothesis 2: different geographic group has the same frequency of sentiment category.

Therefore, we eliminate non-personal posts and use personal posts only to perform the significance test. Besides that, we explore the differences at user and post level separately by using the number of unique users and posts respectively. For gender groups, there are two classes: male and female. Hence, we get two gender groups with sentiment category data each composed of numbers of negative, neutral, and positive posts and another two gender groups of sentiment category data composed of numbers of negative, neutral, and positive users. In Weibo, the platform provides area options of 34 provincial-level administrative areas in China and any place outside China would be unified as overseas. Many people choose not to reveal their real locations for privacy or no specific reasons. For users who do not fill their location in the profile, the location shows as others by default. So, in total, there are 36 areas groups including overseas, and the default setting: others. The area group number 36 is too large, so we decide to regroup the 36 areas into 5 geographical macro-regions: East, Central, West, Special, and Others[10]. Finally, we get 5 groups of sentiment category data composed of numbers of negative, neutral, and positive posts from each region. Another 5 groups of sentiment category data counted similarly but at the user level. We use Chi-square test to testify our null hypotheses of gender and area on sentiment categories. All statistical analyses are performed using scipy.stats library in python.

3. Results

We employ the best-performed model in the evaluation part: the BERT-based alternative meat model to the entire accumulated dataset and get the sentiment analysis results[11].

3.1. Overview of results

To get a statistical overview of the entire dataset, the descriptive statistics data are shown in Table 5. The post numbers and percentages of different gender and post type among 6 classes are displayed in it. We also demonstrate the proportion of each class in the whole dataset and the unique user percentage respectively.

Compared with huge attention from media, personal interest in alternative meat is not that enthusiastic as personal attitudes only compromise 42.1% of the entire posts. For personal posts only, the proportion of negative, neutral, and positive is: 28.77%, 22.91%, 48.32%. At the user level, the proportion of negative, neutral, and positive is: 31.60%, 21.91%, 46.49% which is slightly pessimistic than that at post level. Over half of personal posts hold negative or uncertain opinions towards alternative meat at both situations. We also notice a difference in gender and post type among posts classes. As in Table 5, the rate of publishers being female ranges from the lowest 33.40% in news to the highest 66.28% in positive posts. And the proportion of reposts in personal posts are 66.88%, 52.94%, 45.86% for negative, neutral, and positive posts respectively which are significantly higher compared with those in non-personal posts which are all below 10%. The comparable high number among personal groups shows that individuals are more likely to engage in opinion sharing in terms of personal attitudes towards alternative meat, especially for negative ones.

3.2. Factor analysis results

The frequencies of posts and users per sentiment category differ significantly between male and female, as revealed by the Chi-square test (posts categories: $\chi2 = 560.01$, $p < 0.0001$; users categories: $\chi2 = 581.81$, $p < 0.0001$). The frequencies of posts and users per sentiment category also differ significantly among areas, as revealed by the Chi-square test (posts categories: $\chi2 = 172.56$, $p < 0.0001$; users categories: $\chi2 = 94.16$, $p < 0.0001$). The results show that we may reject the null hypotheses and there is significant difference on the sentiment category proportions among gender and area groups no matter at posts or users level. In other words, area and sentiment category of posts or users are not independent. It can be seen that after user cleaning, the sentiment category distribution differences among gender groups increase slightly while the distribution differences among area groups decrease sharply though still very significant at statistical level according to χ^2 value. The distribution and proportion of post and user sentiment categories among genders and regions are shown in Table 6.

Since the Chi-square test concludes that there is significant association between regions and post or user attitudes, we decide to conduct a post-hoc pairwise comparison to investigate the differences each pairwise using partitions of $\chi2$ method. We reduce the table into multiple 2×3 contingency tables and perform the Chi-square test with applying the Bonferroni corrected alpha level. For posts, the result indicate that all regions differ from each other except the pair of Western China with Eastern China. At user level, the results show that all regions differ from each other except the pair of Eastern China with Others, and the pair of Special with Others.

In Table 6, the proportions of female posts and users are both higher than male group showing higher participation of women in this topic. For female posts, compared with expected value, the observed value of positive class is much higher, but the observed values of neutral and negative classes are much lower while the male posts show opposite tendency. After unique user processing, these tendencies remain same despite the negative proportions of both gender increase slightly.

Five regions show different engagements with Eastern region contributes around half of the entire posts or users. It can be seen that at the user level, Eastern, Special, and Others regions show higher values of negative category and lower values of positive category while Central and Western regions indicate opposite results. The highest levels of positive category and lowest levels of negative category both go to Western and Central region, before and after the unique user cleaning procedures.

We further plot the distributions and percentages of three personal sentiment categories among 36 areas in Fig. 7 at the post level and Fig. 8 at the user level, respectively[12]. Most areas in the front rank of negative levels are from Eastern region or Specials except Tibet and Xinjiang.

3.3. Trends analysis

We illustrate the rolling weekly average number of personal and non-personal posts on the entire data from January 1, 2020, to April 30, 2021, each day in Fig. 9. We can easily get the perception that non-personal posts especially bots and news consist a large part of the whole posts.

We have gathered several important events and news related to alternative meat and products as listed below. We further plot the daily posts number of news, negative, neutral, and positive posts respectively as shown in Fig. 10 and mark the event time in it.

Event 1: Starbucks debuts Beyond Meat products on its new ‘GOOD GOOD’ menu on April 22, 2020, in China.

Event 2: Kentucky Fried Chicken (KFC) Plant-Based Chicken Nuggets debuts in China during April 28–30, 2020.

Event 3: Following the successful trial in June, KFC introduces the plant-based Beyond Burger under the name “New Era series” on October 12, 2020.

Event 4: Beyond Meat announces the launching of a brand-new product called Beyond Pork, developed specifically with Chinese consumers on November 18, 2020.

Event 5: Macdonald’s serves its first plant-based meat option in China from March 24, 2021.

[10] Eastern China consists of Liaoning, Hebei, Beijing, Tianjin, Shandong, Jiangsu, Shanghai, Zhejiang, Fujian, Guangdong, Guangxi and Hainan; Central China encompasses Heilongjiang, Jilin, Inner Mongolia, Shanxi, Henan, Anhui, Hubei, Jiangxi and Hunan; Western China includes Xinjiang, Gansu, Qinghai, Ningxia, Shaanxi, Tibet, Sichuan, Chongqing, Guizhou and Yunnan; Special covers Hong Kong, Marco, Taiwan, and Overseas; Others: others.

[11] Although the posts’ details cannot be shared publicly due to privacy concerns, we provide all the ids and predicted labels of the sample and whole dataset as supplementary materials, sample and results. All the posts could be accessed by their ids through corresponding link https://m.weibo.cn/detail/id.

[12] We use abbreviations for 36 areas. Tibet(XZ), Xinjiang(XJ), Overseas(OS), Tianjing(TJ), Taiwan(TW), Shangdong(SD), Others(OS), Guangdong(GD), Fujian(FJ), Hainan(HI), Beijing(BJ), Shanghai(SH), Jiangsu(JS), Sichuan(SC), Heilongjiang(HL), Yunnan(YN), Zhejiang(ZJ), Shaanxi(SN), Ningxia(NX), Gansu(GS), Shanxi(SX), Hebei(HE), Guangxi(GX), Anhui(AH), Inner Mongoria (IM), Hubei(HB), Qinghai(QH), Chongqing(CQ), Henan(HA), Jilin(JL), Liaoning(LN), Jiangxi(JX), Macao(MO), Hong Kong(HK), Hunan(HN), Guizhou(GZ).

Table 5
Descriptive statistics data.

Variables			Class 0	Class 1	Class 2	Class 3	Class 4	Class 5
Posts Number	Gender	Male	3673	9341	1422	2754	1698	2866
		Female	3743	4685	1326	2307	2333	5634
	Type	Original	7393	13173	2637	1676	1897	4602
		Repost	23	853	111	3385	2134	3898
Posts percentage	Gender	Male	49.53%	66.60%	51.75%	54.42%	41.12%	33.72%
		Female	50.47%	33.40%	48.25%	45.58%	57.88%	66.28%
	Type	Original	99.69%	93.92%	95.96%	33.12%	47.06%	54.14%
		Repost	0.31%	6.08%	4.04%	66.88%	52.94%	45.86%
Class users number			6525	7045	2157	4515	3131	6644
Class posts number			7416	14026	2748	5061	4031	8500
Unique user percentage in Class			89.33%	50.23%	78.49%	89.21%	77.67%	78.16%
Class posts proportion			17.75%	33.57%	6.58%	12.11%	9.65%	20.34%

Table 6
Sentiment category distribution among gender and region groups at post and user level.

Variables		Negative		Neutral		Positive		Total Percentage
		Number	Percentage	Number	Percentage	Number	percentage	
Post	Male	2754(+)	37.63%	1698(+)	23.20%	2866(-)	39.16%	41.60%
	Female	2307(−)	22.45%	2333(−)	22.71%	5634(+)	54.84%	58.40%
User	Male	2363(+)	39.12%	1578(+)	26.12%	2100(−)	34.76%	42.27%
	Female	2152(−)	26.09%	1553(−)	18.83%	4544(+)	55.09%	57.73%
Post	Eastern	2525(−)	28.42%	2182(+)	24.56%	4177(−)	47.02%	50.50%
	Central	491(−)	22.68%	589(+)	27.21%	1085(+)	50.12%	12.31%
	Western	395(−)	26.80%	328(−)	22.25%	751(+)	50.95%	8.38%
	Special	526(+)	37.65%	277(−)	19.83%	594(−)	42.52%	7.94%
	Others	1124(+)	30.61%	655(−)	17.84%	1893(+)	51.55%	20.87%
User	Eastern	2326(+)	31.97%	1659(+)	22.80%	3291(−)	45.23%	50.92%
	Central	444(−)	24.82%	427 (+)	24.87%	918(+)	51.31%	12.52%
	Western	347(−)	28.40%	223(−)	18.25%	652(+)	53.36%	8.55%
	Special	426(+)	37.60%	213(−)	18.80%	494(−)	43.60%	7.93%
	Others	972(+)	33.87%	609(−)	21.22%	1289(−)	44.91%	20.08%

Note: (+) or (−) indicate that the observed value is higher or lower than the expected value according to the global Chi-square test.

Public attention reaches several peaks at certain time points which correspond to some important events in the real world timeline. For news posts, the largest news peak occur on November 18, 2020, the same time as event 5: the announcement of launching Beyond Pork from Beyond Meat. The second flow of peaks occurs from late April to early May 2020 which is in line with Event 3 and 4, the debut of Beyond Meat products with Starbucks and plant-based products from KFC. Rising trends of negative posts in early May 2020 and around October 12, 2020 correspond to Event 2 and 3, trials of plant-based nuggets and burgers from KFC. The second peak of neutral posts comes shortly after Event 4 correspond to the commercial promotions of KFC around late April to early May 2020. Two positive peaks occurring during October 11 to 12, 2020 and April 2021 correspond to Event 3 and 5, the promotion periods of plant-based products from KFC and Macdonald's separately.

We also find that early peaks of news, negative, and neutral posts in January 2020 and neutral posts peak around 25 October 2020 remain unknown and cannot be correlated with any Event listed. We analyze the news posts at that time and find most news can be divided into two categories. One is about an article from Washington Post published on December 26, 2019: 'Dear men: There's no evidence that eating Impossible Whoppers will give you breasts.' Another is an outburst of financial news about alternative meat stock and market. We select representative news and list them as post 1, 2, and 3 in supplementary material, Table S.1. For negative first peak, we select representative posts from negative posts published during the first peak from January 19 to 20, 2020 and list them as post 4 and 5 in supplementary material, Table S.1. We also check through the first peak and third peak of neutral posts. We find that large amount of first peak posts are reposts from an introduction video of alternative meat published by a Key Opinion Leader (KOL) called PaperClip[13], a group which produces educational films. The third peak of neutral posts contains many posts with similar hashtag:'#你会选择植物肉当正餐吗#' (#Will you choose plant-based meat for dinner? #). The representative posts are shown as post 6, 7 in supplementary material, Table S.1.

3.4. Text analysis results

In this section, we focus on the public's main concerns among different class groups specifically. For personal posts in negative, neutral, and positive classes, we introduce the topic model to extract information on public concerns. Topic Model is a probabilistic generative model often applied in information retrieval, machine learning, and natural language processing to find that the hidden topics in discrete data like text corpora. The most classic one is latent dirichlet allocation (LDA) model (Blei, Ng, & Jordan, 2003). LDA is a hierarchical Bayesian model at the three-level of word, topic, and document. For every document in the corpora, a finite mixture over an underlying set of topics characterized by a distribution over the words and the relative importance of the topics is modeled. Stop words, punctuations, usernames, and web links are removed, then the contents of posts are segmented by jieba to get tokens for dictionary and corpus construction next. Table 7 shows the topics discovered by LDA with 5 topics each for three groups.

In Table 7, for negative posts, topics 1, 3, and 5 contain words like:

[13] The account Paperclip was caught up in controversy and disappeared on Weibo in July 2021.

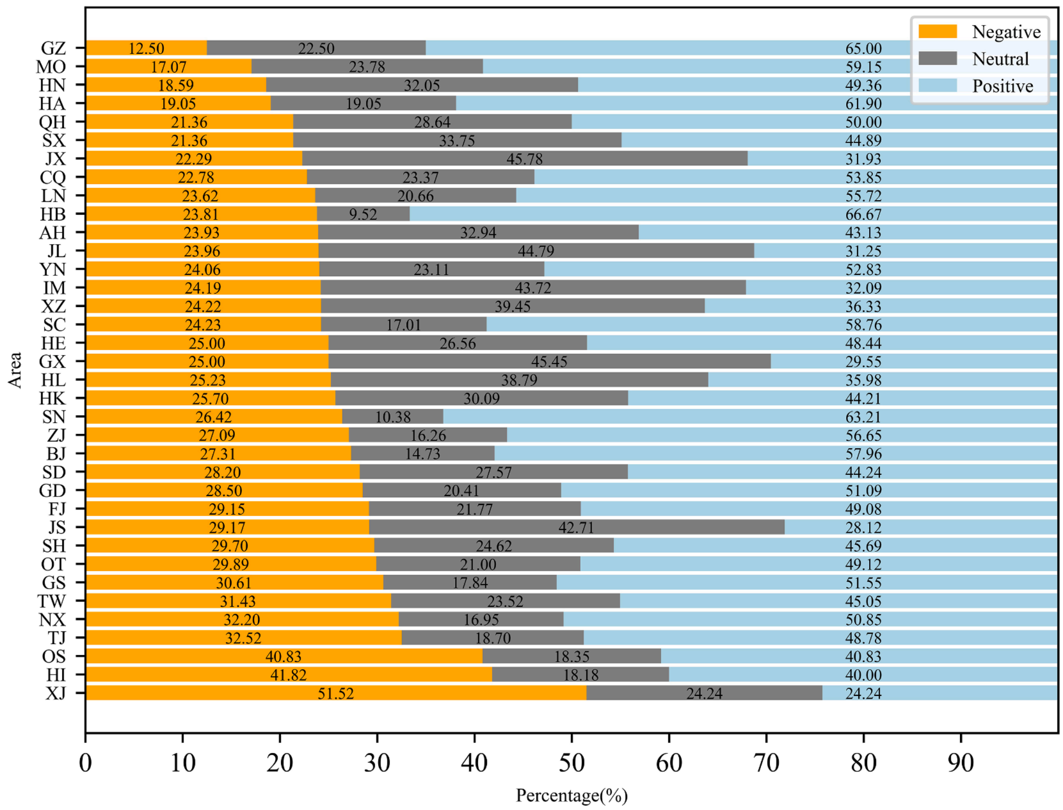

Fig. 7. The distribution and percentage of sentiment category among 36 areas at the post level. The orange, grey and light blue bars from left to right represent the distribution of negative, neutral, and positive attitudes in personal posts with the number showing its percentage respectively. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

genetically modified (GM), truth, America along with the pejorative Chinese slang term for foreigners. Topics 2 and 4 are mainly about alternative meat alike substitutes: tofu and meat flavor mention adding excessive amount of fat. To get a full scope of the content, we select posts containing the word GM, America, *Guizi*, meat flavor, or Tofu from negative class respectively. From the posts that get most reposts, thumbs up, and comments in each word cluster, considering original and repost ones separately, ten representative posts are selected and shown in supplementary material, Table S.2. Generally, strong accusations and condemns towards foreign companies and tycoons for promoting harmful alternative meat to Chinese consumers have been observed among these negative posts. From post 1, 2, 3, 8, and 9, untrusty and hatred towards genetically modified meat are found. These posts indicate beliefs that genetically modified meat (cultured meat) backed by foreign companies and tycoons like Bill Gates and Li Ka-shing is rather harmful and these players in capitals are sinful for promoting such harmful and horrible products to Chinese consumers. From post 7, we also observe the idea that plant-based meat is not healthy as it claims in the propaganda. Driven by profits, foreign companies mislead consumers by adding excessive additives to have meat flavor despite potential health risks. Alongside with post 10, we notice the great emphasis they put on soy-based product like tofu, that tofu resembles much with alternative meat mostly for the same raw materials they use, believing the only trick for alternative meat is adding fat. We see an urge of boycott and ban for alternative meat from posts like post 4, and worries that Chinese nation would be harmed unconsciously if no action taken in post 5. Besides above controversy, we also find negative feedbacks from alternative meat product in post 6 that the products only have meat texture but no real meat flavor and customer's preference for farm-grown meat. We could summarize that many negative opinion posts are devoted to the idea that in order to make profits or other secret reasons, international companies or capitals promote the alternative meat with deceptive ads and unscrupulous propaganda risking Chinese consumers' healthiness.

For neutral, through words (finance, Bill Gates, market, IPO, and Li Ka-shing) from topics 2, 3, and 5, it shows the close relation between tycoons like Bill Gates, Li Ka-shing and the emerging business. For topics 1 and 4, more specific and valuable ideas are obtained as more descriptive words have shown up. We use the same way to find representative posts among all neutral posts and display them in supplementary material, Table S.3. Posts number 1, 2 indicate public curiosity and question towards alternative meat. Posts number 3, 4, 6, and 'expensive or not' from topic 1 show public's curiosities on alternative meat and concerns on the price. People show no negative attitudes as they know little about this new things but there is still a long way before they take any purchase actions because of the potential high price. Astonishments and concerns about safety problem are expressed in post 7 and 8, alongside with words 'estrogen, men, feminization, beans, and view'. It indicates the public attentions on the possible danger brought by consuming alternative meat, like feminization on men by estrogen from beans which are the main ingredients of plant-based meat.

For positive posts, it can be seen that people mainly focus on the specific food promotions as specific brands, slogans, and campaign names have been brought up (new era series, burger, Starbucks) with product feedback (delicious, eat, and chicken). Environment sustainability and the vegetarian diet could also be detected (future, vegetarian diet, and flower). Celebrity existences are also observed as the appearance of the Chinese celebrity LiYuchun, the ambassador of the Starbucks 'GOOD GOOD' campaign. By selecting posts containing words: delicious, Starbucks, vegetarian diet, future, Macdonald's, we select representative posts from posts that get most thumbs up, reposts, and comments and list them in supplementary material, Table S.4.

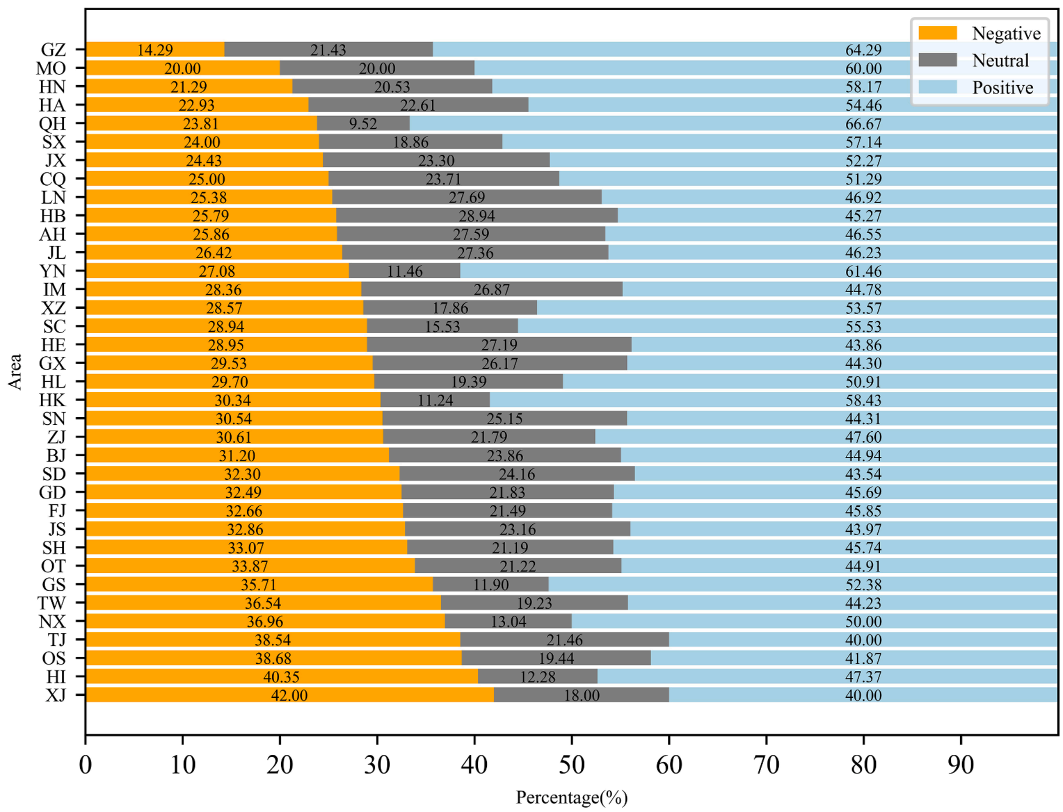

Fig. 8. The distribution and percentage of sentiment category among 36 areas at the user level. The orange, grey and light blue bars from left to right represent the distribution of negative, neutral, and positive attitudes in personal posts with the number showing its percentage respectively. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

4. Discussion

4.1. Acceptance level

We accumulate 41,782 unique alternative-meat-related posts from January 1, 2020, to April 30, 2021, as the entire dataset. With news coverage consists 33.57% of all data, we see enthusiasm from media, government, and industry. About 42.10% of all posts are personal posts expressing negative, neutral, and positive attitudes with a ratio of 12.11%, 9.65%, and 20.34% among all posts respectively. According to our sentiment results, the rate of negative, neutral, and positive posts among personal ones is 28.77%, 22.91%, and 48.32%, respectively. At the user level, with slim increase in negative and decrease in neutral and positive classes the situation is even less optimistic. Despite we get a higher percentage of positive posts than negative ones, over a half posts and users among personal ones show negative or neutral attitudes towards alternative meat.

The results of the previous paper (Dempsey & Bryant, 2020) show that for cultured meat 70% of Chinese consumers show a willingness to try, and 58% are already willing to purchase. Another survey (Liu et al., 2021) indicates about 52.9% of Chinese accept artificial meat as an alternative and 9.6% of them refuse it. More than 70% of Chinese are positive towards cultured meat with a willingness to taste and less than a quarter of the public is opposed to it (Zhang et al., 2020). Considering plant-based meat and cultured meat separately, only 4.4% and 6.7% of Chinese are not at all likely; 33.2% and 33.9% are moderately likely; and 62.4% and 59.3% are extremely likely to buy plant-based meat and cultured meat respectively (Bryant et al., 2019). The acceptance levels of plant-based and cultured meat are both much higher compared with our findings. Only findings from Siegrist and Hartmann (2020) report a similar pattern of public attitudes in China. They find that the acceptance of cultured meat is only 47.5% in China, very close to our result of 48.32%. But the rejection attitude and neutral ones are not explored in that paper. We can conclude that except for results from Siegrist and Hartmann (2020), all previous studies present a much promising market than us. From our sentiment analysis, the positive rate or acceptance of alternative meat is only 48.32% and over a quarter (28.77%) show unwillingness to try and opposition or dissatisfaction towards alternative meat.

What caused the difference between our findings with previous studies remains unknown. But we have several hypotheses for this. First, social media data lack the process of comparison and reasoning and are rather volunteered compared with survey data generated from questionnaires (Wang, Jin, Liu, Li, & Zhang, 2018). For example, as shown above, most posts are linked with important events, news, and product launching. The post could possibly be a quick and flash thought after browsing the news on the phone, leading to the lack of careful consideration and thinking. This assumption could be possible as we have found that compared with neutral and positive posts, negative posts are more spontaneous with a higher proportion of reposting. Also, the voluntary characteristic of social media data that they are published online voluntarily, could also lead to disproportionately visibility from different opinion group if they exhibit different propensities to get involved in such public debate. Previous study has found certain opinion groups would be active more significantly, making their positions are disproportionately visible and impacting the public opinion unequally (Gaisbauer, Pournaki, Banisch, & Olbrich, 2021).

Secondly, as most studies were carried out before 2020, online public views could have changed dynamically over time. On January 7, 2020, the CEO of Impossible Foods made a statement linking the purchase of meat by Chinese consumers with deforestation in Amazon, which kind of

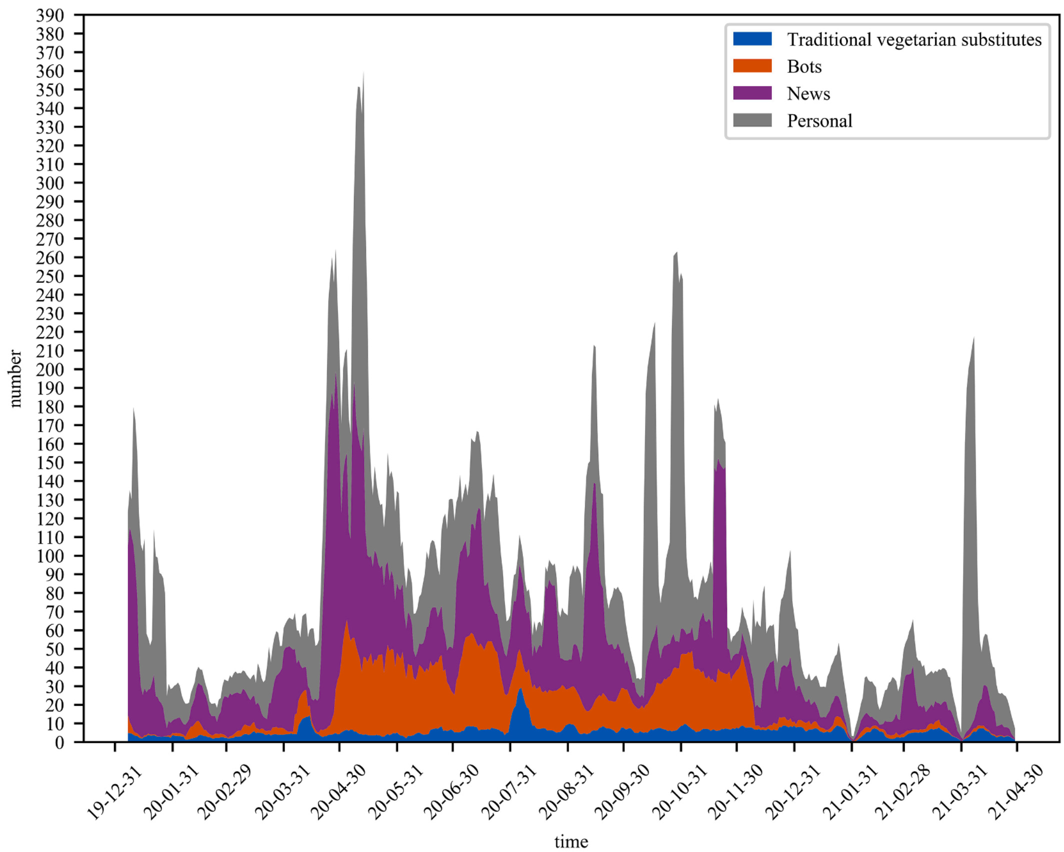

Fig. 9. The stacked rolling weekly average number of personal and non-personal: bots, news, and traditional vegetarian substitutes posts from January 1, 2020 to April 30, 2021.

blames China for consuming too much meat, had caused great backlash in China[14] and sparked strong online accusations that it was "insulting China" among public. As reflected by the 'Amazon smoke' saying in post 5 from supplementary material, Table S.1, the anger initially provoked by the unwelcome declaration could possibly affect the public opinions towards related products.

Last but not least, there are demographical differences between social media users and survey respondents. For example, the percentage of younger users born after 1990 in Weibo is around 78% according to the Weibo's user development report for 2020. Most survey participators bear problems like respondents are all urban citizens from developed areas like Beijing or some are biased towards professional career and higher educational background, but those survey participators are rather balanced in age.

Most previous researches on exploring determinants on attitudes towards accepting alternative meat have found that gender and area both matter. This corresponds to our findings that gender and geography both show significant effects on the public attitudes on alternative meat. Women are more active and tend to have better prospects towards alternative meat. Different regions indicate diverse frequencies of sentiment distribution. Since social media data could be used as volunteered geographic information (VGI) source to show public interest, it can be concluded that Eastern region shows more interest than other places. This could be the results of unbalanced economical level among regions leading different exposure to information. But the less interested places like Western or Central region give better evaluations except Tibet and Xinjiang. The comparable lower level of negative and higher level of positive frequencies from less interested places are perhaps motivated by a lack of general public awareness on the newly product. The special cases of Tibet and Xinjiang could be brought by the lack of representativeness since these are the least engaged areas in China for Weibo users according to the Weibo's user development report for 2020.

4.2. Challenges and obstacles

It is worth noticing that 6.58% of posts of all posts are about traditional vegetarian substitutes and are rather evenly distributed throughout the year. China has long been sculpting and flavoring traditional meat substitutes out of soy, wheat, and mushrooms before any modern novel plant-based meat hit the western market. Previous studies (Dempsey & Bryant, 2020) think that this special food culture in China would help with the embracement of alternative meat. However, from negative posts like post 7 and 10 in supplementary material, Table S.2, we observe frequent mentions of the already existed traditional meat substitutes. They believe alternative meat is nothing more than just adding excessive additive like fat into soy products, depreciating the value of alternative meat. The technology is simplified and the difference is questioned with claims that with various soy-based products available, there is no place left for alternative products which is unhealthy and expensive. Hence, the existence of traditional vegetarian substitutes could be an obstacle instead of a boost for the promotion of alternative meat as previous researchers expected. Possible solutions of those could be differentiation strategies for the alternative meat product to distinguish them from traditional vegetarian substitutes which are widely accepted and cheap.

The controversy towards alternative meat has moved farther away from food itself and became much complex and rather conspiracy

[14] Pat Brown, the chief executive of Impossible Foods said as quoted: 'Every time someone in China eats a piece of meat, a little puff of smoke goes up in the Amazon.' on New York Times, January 7, 2020. Retrieved from: https://www.nytimes.com/2020/01/07/business/fake-pork-china.html

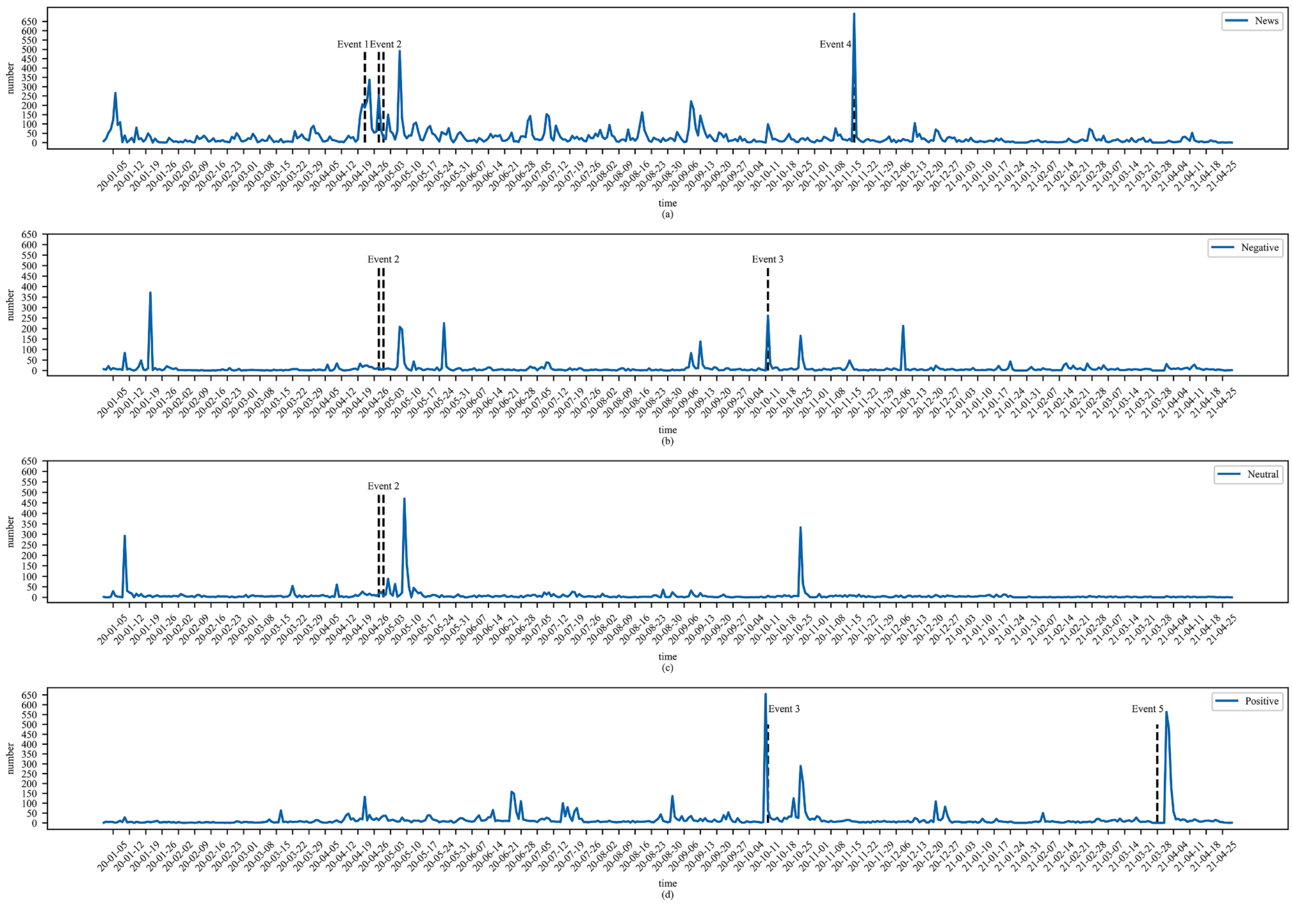

Fig. 10. The daily post number of news (a), negative (b), neutral (c), and positive (d) category respectively during collected period. The vertical dashed lines mark the Events time.

Table 7
Topics from three personal attitudes group.

Class	Topic	Word 1	Word 2	Word 3	Word 4	Word 5
	1	*Guizi* (a pejorative Chinese slang term for foreigners)	Zhifang (fat)	Tianjia (add)	Guoliang (excessive)	Zhongguo (China)
	2	Yourouwei (have meat flavor)	Meirouwei (do not have meat flavor)	Doufu (tofu)	Lu (brine)	Zhifang (fat)
Negative	3	Zhuangjiying (genetically modified)	Xin (new)	Xinnengyuan (new energy)	Shengwu (biology)	Zenme (how)
	4	Zhengxiang (truth)	Haocheng (claim)	Tianjia (add)	Liang (amount)	Cuo (wrong)
	5	Zhongguo (China)	Meiguo (America)	Lai (come)	Zhuanjiying (genetically modified)	Keyi (could)
	1	Fenxiang (share)	Lvzhou (a pic sharing app)	Dongtai (posts)	Guibugui (expensive or not)	Buxiang (do not want to)
	2	Quanshijie (the whole world)	Rongzi (finance)	Xuanze (choose)	Dongtai (posts)	Fenxiang (share)
Neutral	3	Dadao (reach)	Lijiecheng (Ka-shing Li)	Yuji (predict)	Shangshi (Initial Public Offerings:IPO)	Hahaha (hahaha, laughs)
	4	Cijisu (estrogen)	Nanren (men)	Nvxinghua (feminization)	Dadou (beans)	Guangdian (view)
	5	Yimeiyuan (million USD)	Shichang (market)	Rongzi (finance)	Zhenyou (really)	Biergaici (Bill Gates)
	1	Hanbao (burger)	Chi (eat/taste)	Shidai (era)	Xilie (series)	Keji (technology)
	2	Chi (eat/taste)	Hua (flower)	Weilai (future)	Sushi (vegetarian diet)	Womeng (we)
Positive	3	Liyuchun (A celebrity in China)	Xinbake (Starbucks)	Duo (much)	Haochi (delicious)	Xilie (series)
	4	Chi (eat/taste)	Keyi (could)	Hen (pretty)	Kan (see)	Weilai (future)
	5	Chi (eat/taste)	Zhongguo (China)	Haochi (delicious)	Luxianxian (green fairy)	Jirou (chicken)

oriented as we have noticed conspiracy-linked beliefs that sinful foreign capital players promote unhealthy alternative meat in China for profits and other dark secrets with the potential danger of harming Chinese consumers. The term “conspiracy theory” refers to claims that major social and political events and circumstances were caused by conspiracies secretly plotted by a group of powerful actors, often for sinister purpose (Coady, 2006; Douglas et al., 2019; Keeley, 1999). For example, the conspiracy theories of GM foods and vaccinations claim they are unsafe but companies or capitals hide the evidence and promote them to public for commercial profits or dark secrets.

The conspiracy-linked thoughts shared among negative posts are very similar to conspiracy theories related to GM food and vaccinations. They accuse the plant-based meat company typically foreign ones like Impossible Foods of deceptive propaganda intentionally for profits and condemn the evil capitalists hidden behind like Bill Gates and Li Ka-shing of promoting unsafe genetically modified meat (cultured meat) deliberately. These all correspond to the definition of conspiracy theories which are explanations of events or practices, accusing powerful malevolent groups of plotting for their own benefit secretly against the common good (Uscinski, Douglas, & Lewandowsky, 2017). Above all, we can conclude that there is conspiracy theory on alternative meat among negative posts. The conspiracy theory of alternative meat basically claims that foreign companies and tycoons like Bill Gates are sinful actors in promoting unsafe and unhealthy alternative meat with deceptive propaganda for their own benefits at the cost of Chinese nation’s health.

Some papers link opinions of technology or trust in science with alternative meat attitudes (Siegrist & Hartmann, 2020; Slade, 2018), but only one paper addresses the relationship between the public’s conspiratorial ideation and their attitudes towards meat (Wilks et al., 2019). Conspiratorial ideation means a general predisposition to believe that conspiracies happen. This study finds that people with conspiratorial ideation who believe in four popular conspiracies tend to have absolute opposition to cultured meat. The finding of connection between conspiratorial ideation and opposition to alternative neat supports our discoveries on conspiracy thoughts on alternative meat from negative opinion posts as a basis to some degree. All in all, from social media data, we see the conspiracy theory of alternative meat in China. To our best knowledge, this is the first research that identifies the specific existence of conspiracy theory from alternative meat opposition.

Price and food safety are found as the main reasons for holding neutral thinkers back. Their choices are dependent on the price of the alternative meat with concerns about the possible high price of alternative meat as shown in posts from supplementary material, Table S.3. No further movement like try or purchase will actually be done unless they find the price is acceptable. This price-sensitive corresponds to finding that alternative meat would be better accepted with a competitive price (Michel, Hartmann, & Siegrist, 2021). Also, we see a lack of related knowledge about the safety problems and concerns of the possible danger that feminization on men by estrogen from soy ingredients of plant-based meat. The worries about side effects might stop people from trying alternative meat and keep them lingering around have also been discussed in previous studies. This point could be backed up by findings from Liu et al. (2021) that the Chinese concerns more about safety and unnaturalness, than ethical and environmental issues compared with Western customers. Our recommendation to solve this would be more information provision as it has proven to be effective in improving people’s attitudes on alternative meat (Zhang et al., 2020).

Compared with factors like vegetarian diet and environment friendly which are widely accepted in western countries, for Chinese consumers, more focuses are found to be on the taste of the alternative product, and the product campaign. It can be concluded that plenty of positive attitudes are product targeted given the high proportion of product and brand related topics and posts and high-promotion driven positive posts trends. Thus, using health and the environment as selling points may not be the best choice in China. Interestingly, celebrity existences are newly discovered in our study. It is common to use celebrities to attract public attention in the advertisement industry. So the observed presence of celebrity might be incidental which comes along with intensive advertising. Whether the celebrities will promote the products and if so, to which degree will they do remain unknown and could be considered for further research.

We also notice consumers’ unfamiliarity with alternative meat as some posts (like post 5 in supplementary material, Table S.1) mistake plant-based meat for products from Impossible Foods which actually produces cultured meat instead. The reason could be the lack of knowledge and blurred boundary of alternative meat terms in Chinese. Despite the confusion above and other common concerns like price on both types, we have found that different public focus on them from negative posts that plant-based meat receives much pressure from soy-based product and has potential health risks like feminization and excessive additives while cultured meat gets much blame for being genetically modified related. Further work on consumer behavior analysis could target them respectively and alternative meat companies should conduct promotions more specifically to mark themselves off.

5. Limitations

In this paper, we use social media data to investigate Chinese public interest and idea of alternative meat: plant-based meat and cultured meat together. The study method could provide numerous profiles of phenomenon and insights to broaden the scope of survey data for further research as a complementary source. However, as Vidal et al., (2015) say, no method is a panacea. Unavoidably, we encounter the problems that they point out that social media users are not representative with its skewed distribution. Weibo are demographically biased towards young users. Weibo covers most places in China but some developed areas like Beijing, Shanghai and other coastal provinces show higher engagement. However, this systematic sample bias problems could be improved with more diverse active users within time. Also, it may provide us more voices from young-aged consumers who are hard to recruit for survey (Vidal et al., 2015).

Furthermore, compared with surveys, the social media data can only provide limited backgrounds of respondents. We accumulate information like thumbs-up, reposts, comments number, publish tool, number of followers, and number of following but only include the gender and area for factor analysis due to page limit. And the educational and economical background cannot be accessed through our method which are usually important factors in consumer behavior analysis.

Besides that, we consider plant-based meat and cultured meat together for a comprehensive analysis. Although we have found that different opinions and focuses on plant-based meat and cultured meat, we do not obtain the acceptance rate for them respectively due to the mixed usages of terms in online data.

6. Conclusion

This paper investigates public attitudes and possible determinants towards alternative meat in China in an innovative way through social media data by using transfer learning method. We apply a more fine-grained classification method for consumer attitudes analysis by separating personal posts from non-personal ones and dividing posts into 6 categories in total. A large part of posts’ peaks especially news and positive ones are driven by related event and commercial promotions respectively as revealed by the trends. At odds with previous studies on Chinese customers, a less optimistic result is obtained compared with survey findings. Over half of the personal posts are negative or neutral oriented. Gender and area are both found to be influential to the public attitudes towards alternative meat. Despite significant effects are both observed at post and user level, the region differences are lower on the user level while the gender differences opposite. The conspiracy theory is discovered in alternative meat oppositions for the first time. The obstacles brought by strong competitor: traditional vegetarian substitutes like tofu are also discovered in rejection feelings. Worries on price and safety concerns on unknown side effects of alternative meat are confirmed from the hesitation of those in neutrality. Product campaigns are found to be the most contributing part for positive attitudes among the public. Vegetarian diet and environmental sustainability are also noticed among positive feelings.

Generally speaking, compared with survey respondents, in the social media Chinese public tend to have more pessimistic view towards alternative meat food especially on its health risks and conspiracy theories. The rising of conspiracy theory of alternative meat brings challenges for future promotions of alternative meat. The relationship with GM technology promotes the skeptics of alternative meat. More publicity should give to the information provision about safety problems like feminization, GM technology and product promotions which focus on taste and differentiation strategy to distinguish alternative meat from soy-products.

CRediT authorship contribution statement

Yuan Chen: Conceptualization, Methodology, Formal analysis, Software, Validation, Visualization, Writing – original draft, Writing – review & editing. **Zhisheng Zhang:** Supervision, Project administration, Funding acquisition, Writing – review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.foodqual.2022.104530.

References

Abbar, S., Mejova, Y., & Weber, I. (2015). You tweet what you eat: Studying food consumption through twitter. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems* (pp. 3197–3206). https://doi.org/10.1145/2702123.2702153

Abd-Alrazaq, A., Alhuwail, D., Househ, M., Hamdi, M., & Shah, Z. (2020). Top concerns of tweeters during the COVID-19 pandemic: Infoveillance study. *Journal of Medical Internet Research, 22*(4), e19016. https://doi.org/10.2196/19016

Aleksandrowicz, L., Green, R., Joy, E. J. M., Smith, P., Haines, A., & Wiley, A. S. (2016). The impacts of dietary change on greenhouse gas emissions, land use, water use, and health: A systematic review. *PloS One, 11*(11), e0165797. https://doi.org/10.1371/journal.pone.0165797

Blackburn, K. G., Yilmaz, G., & Boyd, R. L. (2018). Food for thought: Exploring how people think and talk about food online. *Appetite, 123*, 390–401. https://doi.org/10.1016/j.appet.2018.01.022

Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent dirichlet allocation. *Journal of Machine Learning Research, 3*, 993–1022. https://doi.org/10.1162/jmlr.2003.3.4-5.993

Bryant, C., & Barnett, J. (2020). Consumer acceptance of cultured meat: An updated review (2018-2020). *Applied Sciences*, *10*(15), Article 5201. https://doi.org/10.3390/app10155201.

Bryant, C., & Sanctorum, H. (2021). Alternative proteins, evolving attitudes: Comparing consumer attitudes to plant-based and cultured meat in Belgium in two consecutive years. *Appetite, 161*, 105161. https://doi.org/10.1016/j.appet.2021.105161

Bryant, C., Szejda, K., Parekh, N., Desphande, V., & Tse, B. (2019). A Survey of Consumer Perceptions of Plant-Based and Clean Meat in the USA, India, and China. *Frontiers in Sustainable Food Systems, 3*, Article 11. https://doi.org/10.3389/fsufs.2019.00011

Carr, J., Decreton, L., Qin, W., Rojas, B., Rossochacki, T., & Wen, Y. (2015). Social media in product development. *Food Quality and Preference, 40*, 354–364. https://doi.org/10.1016/J.FOODQUAL.2014.04.001

Chriki, S., & Hocquette, J. F. (2020). The myth of cultured Meat: A review. *Frontiers in Nutrition, 7*, Article 7. https://doi.org/10.3389/fnut.2020.00007

Coady, D. (Ed.). (2006). *Conspiracy Theories: The Philosophical Debate*. Ashgate Publishing, Ltd.

Cortes, C., & Vapnik, V. (1995). Support-vector networks. *Machine Learning, 20*(3), 273–297. https://doi.org/10.1023/A:1022627411411

Danner, H., & Menapace, L. (2020). Using online comments to explore consumer beliefs regarding organic food in German-speaking countries and the United States. *Food Quality and Preference, 83*, 103912. https://doi.org/10.1016/j.foodqual.2020.103912

Dempsey, C., & Bryant, C. (2020). Cultured meat: Do Chinese consumers have an appetite? *OSF Prepr*, 1–40.

Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2019). Bert: Pre-training of deep bidirectional transformers for language understanding. In *Proceedings of NAACL-HLT* (pp. 4171–4186).

Dorce, L. C., da Silva, M. C., Mauad, J. R. C., de Faria Domingues, C. H., & Borges, J. A. R. (2021). Extending the theory of planned behavior to understand consumer purchase behavior for organic vegetables in Brazil: The role of perceived health benefits, perceived sustainability benefits and perceived price. *Food Quality and Preference, 91*, 104191. https://doi.org/10.1016/j.foodqual.2021.104191

Douglas, K. M., Uscinski, J. E., Sutton, R. M., Cichocka, A., Nefes, T., Ang, C. S., & Deravi, F. (2019). Understanding Conspiracy Theories. *Political Psychology, 40*(S1), 3–35. https://doi.org/10.1111/pops.v40.S110.1111/pops.12568

Eyheramendy, S., Lewis, D. D., & Madigan, D. (2003). On the naive bayes model for text categorization. In *International workshop on artificial intelligence and statistics* (pp. 93-100). PMLR.

Frank, M. R., Cebrian, M., Pickard, G., & Rahwan, I. (2017). Validating Bayesian truth serum in large-scale online human experiments. *PLoS ONE*, *12*(5). Article e0177385. https://doi.org/10.1371/journal.pone.0177385.

Gaisbauer, F., Pournaki, A., Banisch, S., & Olbrich, E. (2021). Ideological differences in engagement in public debate on Twitter. *PLOS ONE*, *16*(3), Article e0249241. https://doi.org/10.1371/JOURNAL.PONE.0249241.

Ghani, N. A., Hamid, S., Targio Hashem, I. A., & Ahmed, E. (2019). Social media big data analytics: A survey. *Computers in Human Behavior, 101*, 417–428. https://doi.org/10.1016/j.chb.2018.08.039

Godfray, H. C. J., Aveyard, P., Garnett, T., Hall, J. W., Key, T. J., Lorimer, J., … Jebb, S. A. (2018). Meat consumption, health, and the environment. *Science, 361* (6399). https://doi.org/10.1126/science:aam5324

Gómez-Luciano, C. A., de Aguiar, L. K., Vriesekoop, F., & Urbano, B. (2019). Consumers' willingness to purchase three alternatives to meat proteins in the United Kingdom, Spain, Brazil and the Dominican Republic. *Food Quality and Preference, 78*, 103732. https://doi.org/10.1016/j.foodqual.2019.103732

Grasso, A. C., Hung, Y., Olthof, M. R., Verbeke, W., & Brouwer, I. A. (2019). Older consumers' readiness to accept alternative, more sustainable protein sources in the European Union. *Nutrients*, *11*(8), Article 1904. https://doi.org/10.3390/nu11081904.

Grebitus, C., & Bruhn, M. (2008). Analyzing semantic networks of pork quality by means of concept mapping. *Food Quality and Preference, 19*(1), 86–96. https://doi.org/10.1016/j.foodqual.2007.07.007

Guntuku, S. C., Sherman, G., Stokes, D. C., Agarwal, A. K., Seltzer, E., Merchant, R. M., & Ungar, L. H. (2020). Tracking Mental Health and Symptom Mentions on Twitter During COVID-19. *Journal of General Internal Medicine, 35*(9), 2798–2800. https://doi.org/10.1007/S11606-020-05988-8

Hamilton, L. M., & Lahne, J. (2020). Fast and automated sensory analysis: Using natural language processing for descriptive lexicon development. *Food Quality and Preference, 83*, 103926. https://doi.org/10.1016/j.foodqual.2020.103926

Hartmann, C., & Siegrist, M. (2017). Consumer perception and behaviour regarding sustainable protein consumption: A systematic review. *Trends in Food Science & Technology, 61*, 11–25. https://doi.org/10.1016/j.tifs.2016.12.006

Heikinheimo, V., Minin, E. D., Tenkanen, H., Hausmann, A., Erkkonen, J., & Toivonen, T. (2017). User-Generated Geographic Information for Visitor Monitoring in a National Park: A Comparison of Social Media Data and Visitor Survey. *ISPRS International Journal of Geo-Information, 6*(3), 85. https://doi.org/10.3390/IJGI6030085

Ioffe, S., & Szegedy, C. (2015, June). Batch normalization: Accelerating deep network training by reducing internal covariate shift. In *International conference on machine learning* (pp. 448-456). PMLR. https://dl.acm.org/doi/10.5555/3045118.3045167.

Jaeger, S. R., & Ares, G. (2017). Dominant meanings of facial emoji: Insights from Chinese consumers and comparison with meanings from internet resources. *Food Quality and Preference, 62*, 275–283. https://doi.org/10.1016/J.FOODQUAL.2017.04.009

Jaeger, S. R., Lee, S. M., Kim, K. O., Chheang, S. L., Jin, D., & Ares, G. (2017). Measurement of product emotions using emoji surveys: Case studies with tasted foods and beverages. *Food Quality and Preference, 62*, 46–59. https://doi.org/10.1016/J.FOODQUAL.2017.05.016

Jaeger, S. R., & Rasmussen, M. A. (2021). Importance of data preparation when analysing written responses to open-ended questions: An empirical assessment and comparison with manual coding. *Food Quality and Preference, 93*, 104270. https://doi.org/10.1016/j.foodqual.2021.104270

Jaeger, S. R., Roigard, C. M., & Ares, G. (2018). Measuring consumers' product associations with emoji and emotion word questionnaires: Case studies with tasted foods and written stimuli. *Food Research International, 111*, 732–747. https://doi.org/10.1016/J.FOODRES.2018.04.010

Jaeger, S. R., Vidal, L., & Ares, G. (2021). Should emoji replace emotion words in questionnaire-based food-related consumer research? *Food Quality and Preference, 92*, 104121. https://doi.org/10.1016/j.foodqual.2020.104121

Jaeger, S. R., Vidal, L., Kam, K., & Ares, G. (2017). Can emoji be used as a direct method to measure emotional associations to food names? Preliminary investigations with consumers in USA and China. *Food Quality and Preference, 56*, 38–48. https://doi.org/10.1016/J.FOODQUAL.2016.09.005

Keeley, B. L. (1999). Of conspiracy theories. *The Journal of Philosophy, 96*(3), 109–126. https://doi.org/10.2307/2564659

Kuttschreuter, M.Ô., Rutsaert, P., Hilverda, F., Regan, Á., Barnett, J., & Verbeke, W. (2014). Seeking information about food-related risks: The contribution of social media. *Food Quality and Preference, 37*, 10–18. https://doi.org/10.1016/J.FOODQUAL.2014.04.006

Laguna, L., Fiszman, S., Puerta, P., Chaya, C., & Tárrega, A. (2020). The impact of COVID-19 lockdown on food priorities. Results from a preliminary study using social media and an online survey with Spanish consumers. *Food Quality and Preference, 86*, 104028. https://doi.org/10.1016/j.foodqual.2020.104028

Liu, J., Hocquette, É., Ellies-Oury, M. P., Chriki, S., & Hocquette, J. F. (2021). Chinese consumers' attitudes and potential acceptance toward artificial meat. *Foods, 10*(2), 353. https://doi.org/10.3390/foods10020353

Martí, P., Serrano-Estrada, L., & Nolasco-Cirugeda, A. (2019). Social Media data: Challenges, opportunities and limitations in urban studies. *Computers, Environment and Urban Systems, 74*, 161–174. https://doi.org/10.1016/J.COMPENVURBSYS.2018.11.001

Michel, F., Hartmann, C., & Siegrist, M. (2021). Consumers' associations, perceptions and acceptance of meat and plant-based meat alternatives. *Food Quality and Preference, 87*, 104063. https://doi.org/10.1016/j.foodqual.2020.104063

Moranges, M., Rouby, C., Plantevit, M., & Bensafi, M. (2021). Explicit and implicit measures of emotions: Data-science might help to account for data complexity and heterogeneity. *Food Quality and Preference, 92*, 104181. https://doi.org/10.1016/j.foodqual.2021.104181

Onwezen, M. C., Bouwman, E. P., Reinders, M. J., & Dagevos, H. (2021). A systematic review on consumer acceptance of alternative proteins: Pulses, algae, insects, plant-based meat alternatives, and cultured meat. *Appetite, 159*, 105058. https://doi.org/10.1016/j.appet.2020.105058

Peschel, A. O., Kazemi, S., Liebichová, M., Sarraf, S. C. M., & Aschemann-Witzel, J. (2019). Consumers' associative networks of plant-based food product communications. *Food Quality and Preference, 75*, 145–156. https://doi.org/10.1016/j.foodqual.2019.02.015

Pindado, E., & Barrena, R. (2020). Using Twitter to explore consumers' sentiments and their social representations towards new food trends. *British Food Journal, 123*(3), 1060–1082. https://doi.org/10.1108/BFJ-03-2020-0192

Puerta, P., Laguna, L., Vidal, L., Ares, G., Fiszman, S., & Tárrega, A. (2020). Co-occurrence networks of Twitter content after manual or automatic processing. A case-study on "gluten-free". *Food Quality and Preference, 86*, 103993. https://doi.org/10.1016/j.foodqual.2020.103993

Samoggia, A., Riedel, B., & Ruggeri, A. (2020). Social media exploration for understanding food product attributes perception: the case of coffee and health with Twitter data. *British Food Journal*, *122*(12), 3815-3835. https://doi.org/10.1108/BFJ-03-2019-0172.

Siegrist, M., & Hartmann, C. (2020). Perceived naturalness, disgust, trust and food neophobia as predictors of cultured meat acceptance in ten countries. *Appetite, 155*, 104814. https://doi.org/10.1016/j.appet.2020.104814

Slade, P. (2018). If you build it, will they eat it? Consumer preferences for plant-based and cultured meat burgers. *Appetite, 125*, 428–437. https://doi.org/10.1016/j.appet.2018.02.030

Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., & Salakhutdinov, R. (2014). Dropout: A Simple Way to Prevent Neural Networks from Overfitting. *Journal of machine learning research*, *15*(1), 1929-1958. https:// dl.acm.org/doi/abs/10.5555/2627435.2670313.

Tian, G., Lu, L., & McIntosh, C. (2021). What factors affect consumers' dining sentiments and their ratings: Evidence from restaurant online review data. *Food Quality and Preference, 88*, 104060. https://doi.org/10.1016/j.foodqual.2020.104060

Tilman, D., & Clark, M. (2014). Global diets link environmental sustainability and human health. *Nature, 515*(7528), 518–522. https://doi.org/10.1038/nature13959

Uscinski, J. E., Douglas, K., & Lewandowsky, S. (2017). Climate Change Conspiracy Theories. *Oxford Research Encyclopedia of Climate Science*. https://doi.org/10.1093/ACREFORE/9780190228620.013.328.

Vidal, L., Ares, G., Blond, M. Le, Jin, D., & Jaeger, S. R. (2020). Emoji in open-ended questions: A novel use in product research with consumers. *Journal of Sensory Studies*, *35*(6), Article e12610. https://doi.org/10.1111/JOSS.12610.

Vidal, L., Ares, G., & Jaeger, S. R. (2016). Use of emoticon and emoji in tweets for food-related emotional expression. *Food Quality and Preference, 49*, 119–128. https://doi.org/10.1016/J.FOODQUAL.2015.12.002

Vidal, L., Ares, G., Machín, L., & Jaeger, S. R. (2015). Using Twitter data for food-related consumer research: A case study on "what people say when tweeting about different eating situations". *Food Quality and Preference, 45*, 58–69. https://doi.org/10.1016/J.FOODQUAL.2015.05.006

Wang, Z., Jin, Y., Liu, Y., Li, D., & Zhang, B. (2018). Comparing Social Media Data and Survey Data in Assessing the Attractiveness of Beijing Olympic Forest Park. *Sustainability, 10*(2), 382. https://doi.org/10.3390/SU10020382

Wilks, M., Phillips, C. J. C., Fielding, K., & Hornsey, M. J. (2019). Testing potential psychological predictors of attitudes towards cultured meat. *Appetite, 136*, 137–145. https://doi.org/10.1016/j.appet.2019.01.027

Zhang, M., Li, L., & Bai, J. (2020). Consumer acceptance of cultured meat in urban areas of three cities in China. *Food Control, 118*, 107390. https://doi.org/10.1016/j.foodcont.2020.107390