ABSTRACT

While COVID-19 has impacted humans for a long time, people search the web for pandemic-related information, causing anxiety. From a theoretic perspective, previous studies have confirmed that the number of COVID-19 cases can cause negative emotions, but how statistics of different dimensions, such as the number of imported cases, the number of local cases, and the number of government-designated lockdown zones, stimulate people’s emotions requires detailed understanding. In order to obtain the views of people on COVID-19, this paper first proposes a deep learning model which classifies texts related to the pandemic from text data with place labels. Next, it conducts a sentiment analysis based on multi-task learning. Finally, it carries out a fixed-effect panel regression with outputs of the sentiment analysis. The performance of the algorithm shows a promising result. The empirical study demonstrates while the number of local cases is positively associated with risk perception, the number of imported cases is negatively associated with confidence levels, which explains why citizens tend to ascribe the protracted pandemic to foreign factors. Besides, this study finds that previous pandemic hits cities recover slowly from the suffering, while local governments’ spending on healthcare can improve the situation. Our study illustrates the reasons for risk perception and confidence based on different sources of statistical information due to cognitive bias. It complements the knowledge related to epidemic information. It also contributes to a framework that combines sentiment analysis using advanced deep learning technology with the empirical regression method.

1 Introduction

Undoubtedly, COVID-19 has become one of severe pandemics all over the world. It has caused hundreds of thousands of infected cases and Thousands of deaths daily[1]. The virus cannot be eliminated in the short term, therefore people will be forced to coexist with the Coronavirus in the post-pandemic era. Under such conditions, small-scale local
epidemics caused by COVID-19 often attract great public attention\[2\]. Former COVID-19 studies which concerns the role of information torch on issues including daily sentiment change monitoring algorithms\[3\]\[4\], mental health toll\[5\]\[2\], infodemic during pandemic\[6\], and impact on education\[7\]. Scholars also explored the sources of information for panic about COVID-19 through questionnaires\[8\]. citizens’ negative emotions put severe pressure on governments, leading to health policies changing drastically in America\[9\].

In this study, we are interested in the impact of coronavirus statistics on the emotional status, that is, individuals’ risk perception of infection and confidence in eradicating the pandemic. A wealth of evidence demonstrates that bad emotions such as anxiety are attributed to an attentional bias as well as an information processing bias that operates to identify threats and risks when facing massive information\[10\]\[11\]. However, the proportions of texts that express anxiety or confidence change differently with the pandemic severity, leading to a sneaking suspicion that various kinds of information play diverse roles in information processing and understanding\[8\]. Understanding roles of sources of information is critical for advancing cognition research and improving the level of policy-making.

Meanwhile, Sentiment analysis, which is also called opinion mining, plays a part in understanding our society, especially in conjunction with social media texts. It aims to discover people’s opinions on specific things such as a product, a company or government policy\[12\]\[13\]. With the widespread use of the internet, everyone is now not only a listener to information but also a communicator of information as they can express their opinion in cyberspace freely. The consequences of text sentiment analysis may be more objective if there is sufficient data. Traditional methods of sentiment analysis on social media text are either dictionary- and rule-based methods\[14\]\[15\]\[16\], or shallow models that focus on carefully designed effective features\[17\], which are tailored to obtain a satisfactory classification of opinion polarities. As for sentiment analysis technology, deep learning methods function better than traditional methods such as the emotional dictionary. Deep neural networks have many different structures and techniques, such as convolutional neural networks (CNNs), recurrent neural networks (RNNs), and attention mechanisms\[18\].

Although existing studies using sentiment analysis have unveiled some interesting phenomena and put forward some valuable arguments, some potential problems still need to be studied. Firstly, some research has demonstrated that information has a significant impact on citizens’ emotions during the pandemic\[8\]\[19\]. But the understanding of why cyberspace turns to be in an uproar is not clear. Former research shows the proportions of positive emotion and negative emotion do not illustrate a significant trend changing with the number of cases, which indicates a complex impact mechanism\[8\]. For instance, there are always intense discussions and polarization about whether the pandemic will end or not among China’s online community. Discovering which kind of information leads to risk perception or confidence will be helpful to have a grasp of it. Next, from methodological perspectives, some research considers texts that are collected with hashtags such as "Coronavirus". These will ignore texts related to other aspects of the disease, such as masks\[20\]. Alternatively, if all social media texts are examined, there will be too much redundancy and noise\[3\]. Additionally, there are many abbreviations and acronyms in Chinese social media texts\[21\]. Previous studies use emotional dictionaries or questionnaires to classify emotions, resulting in a potential bias in results\[8\].

Research demonstrates that perceived risk is related positively to affective responses to risk\[22\]. Attentional bias theory states that people usually focus on different information alternatively when they seek emotional needs\[10\], on account of personal information processing bias\[11\]. A study presents two main factors that cause people to panic: fake news and the number of Coronavirus cases\[8\]. Therefore, we propose the following hypotheses based on the literature mentioned above.

- **H1.a** The number of local cases is positively related to risk perception.
- **H1.b** The number of local cases does not have an effect on confidence levels.

Regarding policy, in some countries and regions, such as mainland China, Hong Kong and Taiwan, the CDC divides cases into two parts: local cases and imported cases, which may cause an attentional bias when residents seek health information. As far as our information goes, imported cases will be quarantined in a fixed location, but they may lead to an endless pandemic\[23\]. Changes in emotions shown in the supplementary material do not seem related. Hence, we set out the following hypotheses:

- **H2.a** The number of imported cases are negatively related to confidence levels.
- **H2.b** The number of imported cases does not have an effect on negatively related to risk perception.

Based on bias cognition theory, this study develops hypotheses and test whether people pay attention to different sources of information when searching for it to estimate personal risk or future state\[10\]. In other words, coronavirus statistics impacts fall into two main categories. Moreover, this study proposes an infusion deep neural network named attention-based channels-LSTM multitask-learning model (ACLMM) to carry out a multi-class sentiment analysis, leading to outcomes suited for regression analysis. In the empirical part, we employ a fixed-effects estimation after model tests, for which the results are rigorous and credible.
2 Background

2.1 Social malaise caused by pandemics

There have been many applications of sentiment analysis and affective computing in regard to pandemics. An early study has shown that people who feel panic are more likely to behave irrationally during a pandemic[24]. Irrational behaviors include expressing stress, escaping from the outbreak site, and challenging official announcements. These behaviors may even lead to recession[25]. A previous study confirmed that outbreaks of disease hurt the mood of internet users[26]. [27] examined worldwide trends of different emotions during the COVID-19 pandemic. [5] found that negative emotions increased, while positive emotions and life satisfaction declined during the first wave of the Coronavirus pandemic. [28] discovered that Coronavirus attracted the attention of a wide range of people at different times in varying intensities. A study determined two main factors that cause people to panic: fake news and the number of Coronavirus cases[8]. However, surprising risk information may change beliefs and improve the accuracy of risk perception[29]. As the pandemic moved into different countries, it affected the countries’ financial markets markedly[30].

Negative emotions cause much harm during pandemics, and researchers can use sentiment analysis to reveal public concerns and panic during the pandemic. The emotions are possible mediating variables in the relationship between COVID-19 pandemic and different fields such as tourism[31], politics[32], education[33], vaccination[34], and finance[35]. To handle hardships, some research stress the importance of governments. High-trust societies which could perform well in pandemic rely on a reliable and professional bureaucracy, a good economic situation, and low population density[36].

In terms of the post-pandemic era, there is so much uncertainty and complexity. Periods after the first pandemic wave are called post-pandemic. A study finds that the proportion of infections belonging to the elderly is particularly small during periods of low prevalence and continuously increases as case numbers increase[37]. The authors model the dynamical adoption of prophylaxis through a two-strategy game and couple it with a SIR spreading model. Some studies discuss the change of world economy and the new challenges of globalization in the post-pandemic era. For instance, [38]’s article proposes that the changes induced by heightened nationalism and protectionism will be marginal rather than fundamental. In contrast, the future world economy will need even more globalization. Transformations are needed in science-policy, economies and governance between post-pandemic[39]. Mental loss has existed for long time after a disaster. Research finds that the level of morbidity of posttraumatic stress disorder four months after a fire disaster remained almost unchanged at 29 months, indicating the long-term nature of posttraumatic stress[40]. College students who self-reported COVID-19 infections were more likely to experience anxiety and depression[41].

2.2 Sentiment analysis

Traditional sentiment analysis is a binary classification task, while affective computing focuses on extracting a set of emotion labels[42]. A direct and traditional technique for text sentiment analysis uses a sentiment analysis dictionary, which calculates the sentiment polarity score with a specific algorithm, such as the naïve Bayes method, after counting the frequency of sentimental words[43]. Conjunction words and adjectives with comparative forms can provide the polarity of emotion, researchers usually choose adjectives as keywords of emotional dictionaries[44][45]. It is also difficult to remove emotional ambiguity according to the context. Therefore, to address the drawbacks mentioned above, some studies use a word vector to construct the dictionary in order to obtain a better word representation[46].

Inspired by biological neurons, an artificial neural network abstracts neuronal computations into a simple mathematical form briefly referred as the "M-P” neuron[47]. A neural network with only one neuron can be regarded as a double-layer neural network with only an input layer and an output layer and no hidden layer. The structure of a neural network can be divided into three parts: an input layer, an output layer and the hidden layers. The output of each neuron can be written as \( f(x) = W_i x_i + b \). A perceptron algorithm describes the way neurons process values. Based on duality theory, this algorithm proposes an iterative method to calculate the weight W and offset b, namely, stochastic gradient descent (SGD). Therefore, the perceptron algorithm can be regarded as a special two-layer neural network. For a long time previously, due to the limitation of computing power, a network with a few layers was considered practical. However, the situation has changed greatly with increases in computing power, which has enabled deep neural networks with many layers and complex architectures to achieve amazing results.

Deep learning methods often perform better than traditional methods in sentiment analysis. To address the drawback of the absence of contextual information, learning word embedding is a way to obtain text features, and word embedding rather than it-idf has been used in many recent works[48][49]. Convolution neural networks (CNNs) belonging to deep learning methods are a kind of special feedforward neural network that were originally used in the field of computer vision. Variants of CNNs used for sentiment analysis applications, such as AC-BiLSTM, handle differences between
sentences through multiple LSTM layers[50]. LSTM networks are a type of recurrent neural network (RNN) capable of learning order dependence in sequence prediction problems[51]. LSTM can be used as an encoder to embed words that act as a one-hot vector, so that the word vector can learn the dependency relationship over a longer distance[52]. To solve the problem of long-distance information loss, a multilayer LSTM or bidirectional LSTM can be constructed, which updates only some of the parameters in each training epoch[53],[54].

Although the schemes mentioned above should be mostly adequate to handle long-distance dependencies in the text, the overall performance of their sentiment analysis is still insufficient. Researchers have proposed two methods for improvement: one is an attention mechanism to learn long-distance information, and the other is multitask learning that learns the language rules. The attention mechanism is inspired by the visual attention mechanism in humans. Humans pay high attention to specific things and reduce attention to other things at the same time. For instance, we can hear our partner’s voice even in a noisy environment.

In a natural language processing task, the attention mechanism makes the neural network know what is important based on the input text (i.e. key-value matching) [55]. As Figure 1 illustrates, $A_w$ represents the weight of attention imposed on sum of weights multiplied by input, which leads to less loss when it is long-history information. The attention mechanism was first applied in machine translation tasks[56]. The attention mechanism has also been widely used in the field of sentiment analysis [57],[58],[?]. In addition, the attention mechanism can also be applied between labels and text matching[59], as well as in aspect-based sentiment analysis[60].

Multitask Learning (MTL) is a learning paradigm in machine learning methods. Its purpose is to exploit the useful information contained in multiple related tasks to help improve the performance of all tasks[61]. MTL can be combined with other learning techniques, including semisupervised learning, supervised learning, unsupervised learning and reinforcement learning. MTL can be used to further improve the performance of a model[62]. Besides, many new models such as Transformer and graph-based representations, have been used in recent years[63],[64].

Figure 1: Attention mechanism

3 Material and methods

In this study, we are interested in how different coronavirus statistics affect public opinion, specifically, risk perception and confidence levels. But a massive amount of unstructured text data hamper our further investigation. To handle the problem, we first propose a deep multi-class sentiment analysis model to find and analyze texts related to the pandemic. Then we conduct a panel data analysis with a fixed-effect model. The study establishes a connection between deep learning methods and econometric methods, the framework of it is illustrated in Figure 2.

3.1 Proposed multi-class sentiment analysis model

In this section, we describe the detailed architecture of ACLMM. CNNs’ characteristics are using kernels to scan the text vector-matrix so that the neural network can learn local information. The local information includes metadata, slang phrases, and emoticons, which is well suited to Chinese social media texts[21]. On the other hand, RNNs, such as long short-term memory (LSTM) and gated recurrent unit networks, can learn global information[65]. Although the
networks are capable of solving the problem of long-distance dependency to some extent, their learning ability will be impaired if the sentences are too long. So we use the attention mechanism to attach importance to keywords. Moreover, the learning process is based on multi-task learning because a single Chinese corpus is small.

The architecture of ACLMM is as following: firstly, converting the text to word vector matrix as input of convolutional layer. Suppose that the shape of the word vector matrix is $W_{s,d}$, where $s$ is the length of text and $d$ is the length of each word vector. The size of the kernel is $k$, the number of channels is $c$. Concatenating different outputs of channels that are in the same position and gaining $S_{s-d+1,c}$. And then the outputs are used as the input of a bi-directional LSTM network in which the length of the hidden unit is $c$. In a specific task, copying the hidden unit of bi-directional LSTM layer as input of a specific LSTM layer that length of the hidden unit is $2c$. The classification result is obtained by attention mechanism and softmax operation on the output after the final unit of the specific LSTM layer. The ACLMM is as follows Figure 3. We give detail information of the model in the supplementary material section 1.

The attention mechanism calculates the output of parallel tasks’ LSTM layer. The calculation formula is as follow:

$$\alpha_t = \frac{\exp(v \cdot h_t)}{\sum_{i=1}^{\infty} \exp(v \cdot h_i)}$$

$$S_{A_w} = \sum_t \alpha_t h_t$$

Where $v$ is the output of the final timestep of LSTM and $\alpha_t$ is the weight of each word. The category of output is based on the softmax function. The details of the model are as following:

3.2 algorithm details

This study uses the CBOW algorithm to train word vectors based on a large-scale corpus that sourced from Natural Language Processing and Information Retrieval sharing platform[1]. We choose the CBOW algorithm because it has a faster training speed. Former research demonstrates that character-level word vector performances better than word-level word vector in several tasks based on Chinese social media[21]. So, we deal with the original corpus according to the character segmentation and train word vectors. Finally, we obtain a word vector matrix with a width of 200-dimension and a length of 6,644(words).

The layer of convolution with multi-channel is used to detect emotion from different aspects. The vectors extracted from the same position but different channels are concatenated to produce a feature map $M = \left[ \overrightarrow{m_0}, \overrightarrow{m_1}, \ldots, \overrightarrow{m_{s-k+1}} \right]$, where $s$ is the length of the text and $k$ is the size of the kernel. The length of $\overrightarrow{m_t}$ is equal to the number of channels.

Bi-directional LSTM consists of cells with cell state and gates mechanism[66]. Cell state $C_t$ depends on the former time cell state $C_{t-1}$ and hidden state $h_{t-1}$, current time input $X_t$. Its calculation formula is as follows:

$$c_t = f_t \odot c_{t-1} + i_t \odot \tanh(W_{xc}x_t + W_{hc}h_{t-1} + b_c)$$

---

1. [http://www.nlpir.org/wordpress/](http://www.nlpir.org/wordpress/)
LSTM adopts the gating mechanism and has three types of gates. All gates are parametric functions that can be trained. First is the input gate, it deals with the former time information and current time input information. The calculation formula is as follows:

$$i_t = \sigma(W_{xi}x_t + W_{hi}h_{t-1} + W_{ci}c_{t-1} + b_i)$$

Forget gate decides how the former time cell influences the current cell’s memory by controlling forget some information. The calculation formula is:

$$f_t = \sigma(W_{xf}x_t + W_{hf}h_{t-1} + W_{cf}c_{t-1} + b_f)$$

Output gate transmits information to the next cell and decides which parts of the cell state should be outputted according to the following equations.

$$o_t = \sigma(W_{xo}x_t + W_{ho}h_{t-1} + W_{co}c_t + b_o)$$

$$h_t = o_t \odot \tanh(c_t)$$

bi-directional LSTM can obtain more information by learning emotional information from two opposite directions. Each task has its layers: an LSTM and an attention layer. The hidden states of the bi-directional LSTM layer are copied to each task’s LSTM layer. As for the bi-directional LSTM layer, the output of a cell from front to back is $h_{front,i}$, it is $h_{back,i}$ from back to front. So the input of a cell in the each task’s uni-directional LSTM is $h_{front,i} \oplus h_{back,i}$. The attention mechanism assigns different weights to words according to different contributions to the emotion. A common way to assign weights to calculate the similarity between hidden state and final output with the vector product is the attention mechanism. In this model, the attention mechanism calculates the output of parallel tasks’ LSTM layer. The calculation formula is as follow:

$$\alpha_t = \frac{\exp(v \cdot h_t)}{\sum_{i=1}^{\infty} \exp(v \cdot h_i)}$$

$$S_{Aw} = \sum_t \alpha_t h_t$$

Where $v$ is the output of the final timestep of LSTM and $\alpha_t$ is the weight of each word. The category of output is based on the softmax function.
Table 1: description of datasets

| Dataset                   | Description                                                                 | Number of categories |
|---------------------------|-----------------------------------------------------------------------------|----------------------|
| Nonline_shopping_10_cats  | More than 60,000 comments data about ten categories of commodities, including positive and negative comments about 30,000 | 2                    |
| Simplifyweibo_4_moods     | There are about 360,000 emotional messages on Sina Weibo, of which 200,000 are joy, and 50,000 are anger, disgust and depression. | 4                    |
| weibo_senti_100k         | More than 100,000 emotional texts on Sina Weibo, with about 50,000 positive and 50,000 negative comments | 2                    |
| Nlpcc2014_sentiment      | More than 10,000 emotional texts, divided into three types: positive, negative and neutralize | 3                    |
| COVID-19_sentiment       | There are over 20,000 social media texts with seven emotional attributes about COVID-19 after we integrate some neutral and ambivalent texts as an uncertain type. Categories are fear, disgust, joy, surprise, confidence, sadness, anger, and uncertainty | 7                    |
| COVID-19_identify        | This study collected and labeled around 9,000 texts from Sina Weibo, half of them were related to COVID-19, and half of them were not. | 2                    |

3.3 Training datasets

Six Chinese datasets about sentiment analysis are used in experiments, and the description of datasets is shown in the Table 1. The first five datasets are derived from previous research works or open-source platforms. This work labels the last dataset.

In this part, we use there criteria namely accuracy ($A_{cc}$), precision ($P_r$), and recall ($R_e$) to assess the performance of the algorithm. Accuracy is used to measure whole classification performance. We use both of them to measure performance in each single class because Precision and recall are contradictory in some cases. These criteria are extensively used in text classification and sentiment analysis tasks. They are calculated as follows:

$$Acc = \frac{TP + TN}{TP + FP + TN + FN}$$

$$Pr = \frac{TP}{TP + FP}$$

$$Re = \frac{TP}{TP + FN}$$

$$F_1 = \frac{2 \times Pr \times Re}{Pr + Re}$$

where $TP$, $TN$, $FP$, and $FN$ are true positive, true negative, false positive, and false negative, respectively.

3.4 Model outputs

The supplementary material (section 1) goes into detail about training datasets, training process, and evaluation. The result based on annotated COVID-19_sentiment dataset compared with previous studies shown in Table 2 and the training result shown in Figure 4, the ACLMM achieved better performance in the dataset. Compared with AC-BiLSTM [50] that has multiple LSTM layers and attention layers, We find it is substantial to use the kernel structure in the task of Chinese informal text sentiment classification, which is in consistence with former research [21]. The MTL has significantly improved the performance of the task. Here we gain some outputs in the following formulas:

$$attention_{it} = \frac{A_{it}}{W_{it}}$$
Figure 4: Accuracy and loss changes in training

\[
\text{emotion}_{it} = \frac{E_{it}}{A_{it}}
\]

where \(A_{it}\), \(W_{it}\) means the total number of texts and the number of texts belonging to the type of attention in day \(t\) from city \(i\), respectively. \(\text{emotion}_{it}\) represents the number of a specific (e.g., fear) texts in day \(t\) from city \(i\). We call \(\text{attention}_{it}\) as pandemic texts, whose accuracy of the task that classifies pandemic texts is 86.02%. Similarly, the precision and recall score of \(\text{emotion}_{it}\) are shown in Table 3. There are seven types of emotion, and it performs pretty well in the class of fear and confidence.

Table 2: Results in different models

| Name                | Accuracy |
|---------------------|----------|
| LSTM [31]           | 0.78     |
| Text-CNN [68]       | 0.7692   |
| bi-LSTM             | 0.8011   |
| AC-BiLSTM [50]      | 0.8210   |
| ACLMM_without_MTL   | 0.8137   |
| ACLMM               | 0.8418   |

3.5 Empirical methodology

Our analysis uses the data of 10 large cities in China over 31 days (Aug 1 - Aug 31), resulting in 462 observations. Since the data consists of a panel of 10 cities for 31 days, where \(N=10\) is much less than \(T=31\), the GMM estimator is not appropriate for our analysis. Definitions of variables are furnished in Table 4. There are three independent variables: attention, fear, and confidence. All of the above variables are calculated according to their proportion in the studied texts based on the chosen categories. Attention is the proportion of COVID-19-related texts to the total number of texts in a city’s social media communications in one day. Fear and confidence are the proportions of fear-related texts and confidence-related texts in a city’s COVID-19-related texts on one day.

control variables: We also try to analysis the historic impacts of each city, so we adopt some control variables to get rid of covariates. A limited government indicates that a government is limited in power. In contrast, A larger government has more government capacity and, correspondingly, more fiscal spending. We choose two indicators as

Table 3: Comparison of the results in each emotion

| Class      | Precision | Recall | F1   |
|------------|-----------|--------|------|
| Fear       | 0.8987    | 0.9405 | 0.9191 |
| Disgust    | 0.6895    | 0.6102 | 0.6475 |
| joy        | 0.8427    | 0.7845 | 0.8126 |
| surprise   | 0.8788    | 0.8985 | 0.8885 |
| confidence | 0.9570    | 0.9727 | 0.9648 |
| Sadness    | 0.7197    | 0.8123 | 0.7632 |
| Anger      | 0.7947    | 0.7579 | 0.7759 |
| Uncertainty| 0.8251    | 0.7992 | 0.8119 |
Table 4: Explanation of the variables

| Variable name     | Explanation                                                                 | Data source                                      | Unit     |
|-------------------|-----------------------------------------------------------------------------|--------------------------------------------------|----------|
| Attention         | the proportion of texts in relation to pandemic                            | output of ACLMM                                  | —        |
| fear              | the proportion of texts which show personal Concerns                         | output of ACLMM                                  | —        |
| confidence        | the proportion of texts which show confidence on the pandemic               | output of ACLMM                                  | —        |
| population outflow| net outflow of the population daily                                         | Baidu Population Migration Data Platform         | person   |
| cases             | The number of local new patients                                           | Government releases report                        | person   |
| Foreigners        | The number of patients from abroad                                          | Government releases report                        | person   |
| Risks             | The number of medium-risk areas and high-risk areas                         | Government releases report                        | —        |
| Distance          | Distance between the current city and nearest one with patients             | Google Maps                                      | km       |
| Pmedical          | The per capita cost of health care for urban residents                       | Chinese Yuan/per person                          |          |
| Pgovernment       | The per capita cost of public service expenditures                          | Wind-Economic Database                            |          |
| Density           | Urban population density                                                    | Wind-Economic Database                            | population/sq.km. |

the government’s capacity: the number of urban health investments per capita and the amount of urban public service expenditures per capita. The spread of virus may be correlated with population density, which will also be a factor affecting people’s emotions, so we choose urban population density as the control variable, which is determined from the wind database. The description of variables is provided in Table 4.

First, we conduct a series of preliminary tests to choose a variety of regression model that is well-suited to our study\[^{69}\]. An F-test is conducted to determine whether a pool model should be used. The result rejects the null hypothesis of the pool model($F = 3.5477, p < 3.285e-09$). Considering exogeneity that includes the different properties of each city, we choose a panel model is appropriate. Wald tests are employed to examine heteroskedasticity and autocorrelation in the data\[^{70}\] [\[^{71}\]]. The Wald test cannot reject the null hypothesis that there is no heteroskedasticity or autocorrelation in the data. Then the Breusch-Pagan LM test illustrates a simultaneous correlation between groups($z = 9.3086, p - value < 2.2e - 16$)\[^{70}\]. Therefore, we use the robust covariance matrix vcovHC (also known as the sandwich estimator) to employ the Hausman test because of heteroskedasticity and autocorrelation\[^{72}\] [\[^{73}\]](chisq = 11.098, p - value = 0.02548). Consequently, we use the fixed-effects panel model. An equation is as follows:

$$y_{ct} = \beta_0 + \beta_1 \ln(cases_{t-1}) + \beta_2 \ln(forign_{t-1}) + \beta_3 \ln(risk_{t-1}) + \beta_4 distance_{t-1} + \beta_5 Pmedical_i + \beta_6 Pgovernment_i + \beta_7 density_i + \gamma_i D_i + \delta t_i + \epsilon_{it}$$

Where $y_{ct}$ representing a category of c(i.e. fear or confidence) on day $t$. These variables are also described in Table 4.

In addition, $D_i$ is the between-group difference of city $i$, $t_i$ is a time trend term, $\epsilon_{it}$ is residual term.

4 Results

4.1 Combining empirical analysis results

Table 5 shows the number of new cases in one city that attracts residents’ attention. It increases residents’ panic (coefficient = 0.405) but has no significant influence on their confidence, while imported cases make citizens lose confidence in eradicating the disease (coefficient = -0.326). However, imported cases have no effect on people’s fright. We cannot reject the null hypothesis H1-2.

In terms of healthcare and city properties, it is generally accepted that spending on healthcare will improve the power to prevent coronavirus; in other words, prevention is better than cure. Our study finds that local governments’ spending on
Table 5: Regression analysis results

| Variable Y       | Model 1 (DV: Fear) | Model 2 (DV: Confidence) | Model 3 (DV: Attention) | Model 4 (DV: Population outflow)(robust test) |
|------------------|--------------------|--------------------------|-------------------------|---------------------------------------------|
| $Cases_{i,t-1}$  | 0.405***           | 0.016                    | 0.788***                | 0.013**                                     |
|                  | (0.103)            | (0.094)                  | (0.130)                 | (0.006)                                     |
| $Foreign_{i,t-1}$| 0.066              | -0.326***                | 0.341***                | -0.004                                      |
|                  | (0.124)            | (0.108)                  | (0.093)                 | (0.006)                                     |
| $Risk_{i,t-1}$   | 0.136**            | 0.031                    | 0.412***                | -0.003                                      |
|                  | (0.067)            | (0.055)                  | (0.103)                 | (0.002)                                     |
| $Distance_{i,t-1}$| 0.013              | 0.003                    | 0.016                   | -0.001                                      |
|                  | (0.020)            | (0.019)                  | (0.023)                 | (0.001)                                     |
| $P_{medical_i}$  | -0.008*            | 0.013***                 | 0.005                   | -0.002***                                   |
|                  | (0.004)            | (0.004)                  | (0.004)                 | (0.000)                                     |
| $P_{government_i}$| 0.001              | -0.012**                 | -0.002                  | 0.002***                                    |
|                  | (0.004)            | (0.005)                  | (0.004)                 | (0.000)                                     |
| $density_i$      | -2.022             | -7.028**                 | -3.018                  | 0.272                                       |
|                  | (3.078)            | (3.220)                  | (2.507)                 | (0.272)                                     |

1 Standard errors are in parenthesis
2 *** p<0.01, ** p<0.05, * p<0.1

Table 6: Intercept term

| Variable Y | Model 1 (DV: Fear) | Model 2 (DV: Confidence) | Model 3 (DV: Attention) | Model 4 (DV: Population outflow)(robust test) |
|------------|--------------------|--------------------------|-------------------------|---------------------------------------------|
| Beijing    | -0.573             | 11.210**                 | 1.610                   | -1.812***                                   |
|            | (3.456)            | (4.358)                  | (3.869)                 | (0.374)                                     |
| Chengdu    | -6.860**           | 9.251*                   | -2.594                  | 1.130***                                    |
|            | (2.775)            | (5.241)                  | (3.676)                 | (0.405)                                     |
| Xi’an      | -4.750*            | -10.527**                | 3.024                   | 0.883***                                    |
|            | (2.845)            | (4.196)                  | (2.448)                 | (0.242)                                     |
| Nanjing    | -3.511             | -6.181**                 | 15.003***               | 0.188                                       |
|            | (2.720)            | (2.676)                  | (3.368)                 | (0.263)                                     |
| Shanghai   | 1.406              | 12.987*                  | 13.550**                | 0.307                                       |
|            | (6.937)            | (7.535)                  | (5.771)                 | (0.601)                                     |
| Zhengzhou  | 9.745              | -8.734                   | 22.730***               | 1.805**                                     |
|            | (7.425)            | (9.972)                  | (7.751)                 | (0.874)                                     |
| Wuhan      | 8.778**            | 8.722                    | 43.851***               | -3.766***                                   |
|            | (4.372)            | (5.532)                  | (4.057)                 | (0.493)                                     |

# observations 462 462 462 462
# subjects 10 10 10 10

1 Standard errors are in parenthesis
2 *** p<0.01, ** p<0.05, * p<0.1
healthcare prevents a panic attack and improves people’s confidence in tackling the pandemic. Specifically, the higher the per capita health cost is, the higher the confidence level.

Finally, we find people trying to escape from where the disease occurred. Conducting a robustness test whose dependent variable is population outflow to avoid a regression fallacy (see supplementary material section 2), we find a significant positive correlation between the net outflow of population and the number of local cases, which is compatible with one research under American circumstances\cite{74}.

4.2 Robust test

we conduct a robustness test to avoid a regression fallacy (see appendix section 2.A). By definition, an economic model is a simplified mathematical representation of a complex interaction of economic variables, as such is built upon certain assumptions. Researchers examine how certain “core” regression coefficient estimates behave when the regression specification is modified by adding or removing regressors. The “fragility” of estimated coefficients indicates that there may be hidden errors in the model\cite{75}.

Former study demonstrates that there is a significant positive correlation between the net outflow of population in a city and the degree of social malaise\cite{74}. The degree of panic caused by the epidemic is in proportion to the escape of urban dwellers (that is, the number of population outflow). It reflects the terror of urban citizens to a certain extent. According to this, our research uses the net outflow of the population as a substitute variable to supersede the variable: fear, which functions as the explained variable. Specifically, the data is obtained by subtracting the population outflow from the inflow.

The regression results show in\cite{7} which depicts the regression results of the robustness test. We find there has been a long-term mental health toll in Wuhan, the city where COVID-19 first broke out in China because people experiencing panic during an epidemic are liable to act irrationally; this phenomenon is also called a social response\cite{24}. Hence, we attempt to discover whether people are still escaping from the sites where disease outbreaks are occurring in the post-pandemic era. This information may help to prevent the triggering of social responses. The results of this study will also provide guidance to public health authorities to better communicate with people and provide public health responses to those who are most frightened by the virus. From the robustness test section in 5.5 and Table 7, we find that people will attempt to escape from the outbreak site in the second wave of COVID-19, and that the social response between the pandemic era and the post-pandemic era is not much different.
Table 7: Regression analysis result of net outflow of the urban population

|          | (1) netout | (2) netout | (3) netout | (4) netout | (5) netout | (6) netout |
|----------|------------|------------|------------|------------|------------|------------|
| L.cases  | 0.013**    | 0.013**    | 0.013**    | 0.012**    | 0.012**    | 0.014**    |
|          | (0.006)    | (0.006)    | (0.006)    | (0.006)    | (0.005)    | (0.005)    |
| L.foreign| -0.004     | -0.004     | -0.004     | -0.005     | -0.006     |            |
|          | (0.006)    | (0.006)    | (0.006)    | (0.006)    | (0.005)    |            |
| L.risk   | -0.003     | -0.003     | -0.003     | -0.002     |            |            |
|          | (0.002)    | (0.002)    | (0.002)    | (0.002)    |            |            |
| L.distance| -0.001     | -0.001     | -0.001     |            |            |            |
|          | (0.001)    | (0.001)    | (0.001)    |            |            |            |
| pmedical | 0.002***   | -0.003     |            |            |            |            |
|          | (0.001)    | (0.005)    |            |            |            |            |
| density  | -0.354     |            |            |            |            |            |
|          | (0.360)    |            |            |            |            |            |
| Chengdu  | 1.130***   | -0.424     | 0.331**    | 0.322**    | 0.312*     | 0.355**    |
|          | (0.405)    | (1.210)    | (0.159)    | (0.163)    | (0.164)    | (0.150)    |
| Xi’an    | 0.883***   | -0.848     | 0.142      | 0.133      | 0.121      | 0.179      |
|          | (0.242)    | (1.577)    | (0.169)    | (0.174)    | (0.175)    | (0.163)    |
| Nanjing  | 0.188      | -0.025     | 0.145      | 0.126      | 0.098      | 0.156      |
|          | (0.263)    | (0.470)    | (0.231)    | (0.235)    | (0.235)    | (0.226)    |
| Shanghai | 0.307      | 2.111      | 0.379***   | 0.376***   | 0.368***   | 0.395***   |
|          | (0.601)    | (2.419)    | (0.128)    | (0.128)    | (0.129)    | (0.118)    |
| Zhengzhou| 1.805**    | -1.315     | 0.148      | 0.124      | 0.087      | 0.138      |
|          | (0.874)    | (2.351)    | (0.251)    | (0.258)    | (0.261)    | (0.255)    |
| Wuhan    | -3.766***  | -3.694***  | -4.146***  | -4.151***  | -4.166***  | -4.112***  |
|          | (0.493)    | (0.559)    | (0.297)    | (0.300)    | (0.303)    | (0.291)    |
| Beijing  | -1.812***  | 2.499      | -0.307     | -0.319     | -0.331     | -0.283     |
|          | (0.374)    | (4.098)    | (0.200)    | (0.205)    | (0.208)    | (0.189)    |
| Shenzhen | 1.666      | 0.475***   | 0.462***   | 0.449***   | 0.508***   |            |
|          | (1.693)    | (1.56)     | (1.159)    | (1.160)    | (1.133)    |            |
| Chongqing|            | -0.169     | -0.179     | -0.191     | -0.130     |            |
|          |            | (0.244)    | (0.250)    | (0.252)    | (0.247)    |            |
| date     | -0.004     | -0.004     | -0.004     | -0.004*    | -0.004**   | -0.002     |
|          | (0.003)    | (0.003)    | (0.003)    | (0.002)    | (0.002)    | (0.002)    |
| _cons    | -1.586***  | 3.401      | 0.000      | 0.000      | 0.015      | -0.072     |
|          | (0.470)    | (5.161)    | (0.250)    | (0.257)    | (0.259)    | (0.255)    |
| Obs.     | 300        | 300        | 300        | 300        | 300        | 300        |

Standard errors are in parenthesis

*** p<0.01, ** p<0.05, * p<0.1
5 Discussion

The availability of abundant textual data from social media can be utilized to recognize people’s emotional changes in the pandemic in a better way. As presented in the Table 5 and Table 6, we can see that among these cities with Coronavirus cases, the risk perception and confidence change differently. The objective our study is to discover which kinds of information lead to risk perception or confidence. We performed a sentiment analysis based on more than 1.6 million text data. We conduct a fixed-effect regression using the analysis outcome and econometric methods.

From a theoretical perspective, the findings echo and enrich cognitive theories. Prominent cognitive theories postulate that an attentional bias toward threatening information contributes to exacerbation of fear and anxiety[76]. Our study demonstrates people tend to seek exclusive sources of information to meet self-risk evaluation. For example, they attribute self-risk to local cases while ascribing a decline in confidence to imported cases. It might explain discrimination is produced due to the cognitive bias that exists in the information processing; therefore, some people blame foreign conspiracies for the protracted pandemic[77, 78].

From a methodological perspective, our research not only leverages a wealth of text data but also proposes a fusion framework combining deep multi-class sentiment analysis and regression analysis. To accurately cope with social media texts related to the pandemic, we need an efficient tool to solve the massive unstructured data. Our proposed sentiment analysis model achieved better performance than former studies. The model is also well suitable for empirical research. The model can have multiple types of outputs. It can narrow the scope of texts to a precise range. The accuracy of our model is 84.58%, which is 7.62% higher than that of Text-CNN[68] and 2.08% higher than that of AC-BiLSTM[50]. Next, Although combining sentiment analysis with econometric methods will contribute to reliable conclusions, a few studies have focused on model tests before an empirical analysis[46]. Our work gives a good framework for it.

In addition, we compare the cities’ differences in light of the results in Table 6, resulting that precisely pandemic hit cities recover slowly from the suffering, which confirms the conclusion of the previous questionnaire survey: there is a long-lasting mental health toll after disasters[40]. This finding also confirms the riffle effect[79]. As Table 6 shows (coefficient = 8.778 * t = 4.327), residents in Wuhan, the city where the first COVID-19 breakout in China was reported, still pay great attention to it. citizens are prone to excessive anxiety and even irrational behaviors under this high-stress circumstance; this phenomenon is also called a social response[24]. As for population outflow, citizens escape from where the disease occurred. Hence, we suggest local governments use information technology to understand where these people are going rather than rigid physical restrictions. Local governments’ spending on healthcare prevents a panic attack and improves people’s confidence in tackling the pandemic. One possible explanation is that governments’ spending on healthcare is a sign of whether they take COVID-19 seriously and are ready for it in any case. It could explain why we observe that large cities have a better ability to respond to COVID-19. We believe that it is more worthwhile for the government to increase investment in the medical field to reduce anxiety. Public health authorities should communicate with people and provide public health responses to those who are frightened by the virus.

6 Conclusion and limitations

In this study, we used a Python scraper to collect more than 1.6 million microblogs. Applying the ACLMM to these data, we unveil a cognitive information bias related to Coronavirus in the post-pandemic era through an econometrics method. With the coronavirus pandemic still ongoing, this work helps governments and researchers predict how public opinion will change in the next wave of the pandemic and have a better understanding of cognitive bias. Taking measures to reduce the adverse effects of this pervasive disease on people’s emotions, especially considering the mental health toll and discrimination against people entering a country in the post-pandemic era[77]. From a methodological perspective, we propose a promising framework combining sentiment analysis using advanced deep learning technology with the empirical regression method, which can be applied in many fields such as social computation.

There are some shortcomings in the research. First, it only concerns Chinese Mainland circumstances and collects Chinese social media texts. Coronavirus statistics of Hong Kong and Taiwan is also divided into local cases and imported cases. Further, we label ambivalent texts as uncertain, leading to the sentiment analysis not being well enough. We consider integrating the aspect-level sentiment analysis method. Besides, there may be additional predictors of risk perception beyond local cases, imported cases, health costs and risk areas that should be examined. In the future, we will expand our research scope, such as considering a non-linear relationship between cases and distance results from the riffle effect[79, 80].


References

[1] World Health Organization, “Who coronavirus (covid-19) dashboard.” [https://covid19.who.int/], 2021.

[2] B. Zhong, Y. Huang, and Q. Liu, “Mental health toll from the coronavirus: Social media usage reveals Wuhan residents’ depression and secondary trauma in the covid-19 outbreak,” *Computers in Human Behavior*, vol. 114, p. 106524, 2021.

[3] X. Lyu, Z. Chen, D. Wu, and W. Wang, “Sentiment analysis on Chinese Weibo regarding COVID-19,” in *CCF International Conference on Natural Language Processing and Chinese Computing*. Springer, 2020, pp. 710–721.

[4] G. Barkur and G. B. K. Vibha, “Sentiment analysis of nationwide lockdown due to COVID-19 outbreak: Evidence from India,” *Asian Journal of Psychiatry*, vol. 51, p. 102089, 2020.

[5] S. Li, Y. Wang, J. Xue, N. Zhao, and T. Zhu, “The impact of COVID-19 epidemic declaration on psychological consequences: A study on active Weibo users,” *International Journal of Environmental Research and Public Health*, vol. 17, no. 6, p. 2032, 2020.

[6] W. Pian, J. Chi, and F. Ma, “The causes, impacts and countermeasures of COVID-19 "infodemic": A systematic review using narrative synthesis,” *Information Processing & Management*, p. 102713, 2021.

[7] S. I. Hofer, N. Nistor, and C. Scheibenzuber, “Online teaching and learning in higher education: Lessons learned in crisis situations,” *Computers in Human Behavior*, vol. 121, p. 106789, 2021.

[8] A. R. Ahmad and H. R. Murad, “The impact of social media on panic during the COVID-19 pandemic in Iraqi Kurdistan: An online questionnaire study,” *Journal of Medical Internet Research*, vol. 22, no. 5, e19556, 2020.

[9] Z. Jin, Z. Peng, T. Vaidhya, B. Schoekopf, and R. Mihalcea, “Mining the cause of political decision-making from social media: A case study of COVID-19 policies across the US states,” in *Findings of the Association for Computational Linguistics: EMNLP 2021*, 2021, pp. 288–301.

[10] A. Mathews and C. MacLeod, “Cognitive approaches to emotion and emotional disorders,” *Annual Review of Psychology*, vol. 45, p. 25, 1994.

[11] S. Hayes and C. R. Hirsch, “Information processing biases in generalized anxiety disorder,” *Psychiatry*, vol. 6, no. 5, pp. 176–182, 2007.

[12] S. Das and M. Chen, “Yahoo! for Amazon: Extracting market sentiment from stock message boards,” in *Proceedings of the Asia Pacific Finance Association Annual Conference (APFA)*, vol. 35, Bangkok, Thailand, 2001, p. 43.

[13] P. Ouyang, J.-J. Wang, and A.-C. J. Chang, “Patients need emotional support: Managing physician disclosure information to attract more patients,” *International Journal of Medical Informatics*, vol. 158, p. 104674, 2022.

[14] S. Rosenthal, N. Farra, and P. Nakov, “Semeval-2017 task 4: Sentiment analysis in Twitter,” in *Proceedings of the 11th International Workshop on Semantic Evaluation (SemEval-2017)*, 2017, pp. 502–518.

[15] J. Xiong, X. Feng, and Z. Tang, “Understanding user-to-user interaction on government microblogs: An exponential random graph model with the homophily and emotional effect,” *Information Processing & Management*, vol. 57, no. 4, p. 102229, 2020.

[16] L. Guo, Y. Li, and D. Sheng, “Modeling and simulating online panic in an epidemic complexity system: An agent-based approach,” *Complexity*, vol. 2021, 2021.

[17] I. Chaturvedi, E. Ragusa, P. Gastaldo, R. Zunino, and E. Cambria, “Bayesian network based extreme learning machine for subjectivity detection,” *Journal of The Franklin Institute*, vol. 355, no. 4, pp. 1780–1797, 2018.

[18] A. Yadav and D. K. Vishwakarma, “Sentiment analysis using deep learning architectures: A review,” *Artificial Intelligence Review*, vol. 53, no. 6, pp. 4335–4385, 2020.

[19] J. Wang, Y. Fan, J. Palacios, Y. Chai, N. Guetta-Jeanraud, N. Obradovich, C. Zhou, and S. Zheng, “Global evidence of expressed sentiment alterations during the COVID-19 pandemic,” *Nature Human Behaviour*, vol. 6, no. 3, pp. 349–358, 2022.

[20] E. Kalampokis, E. Tambouris, and K. Tarabanis, “Understanding the predictive power of social media,” *Internet Research*, 2013.

[21] D. Xu, Z. Tian, R. Lai, X. Kong, Z. Tan, and W. Shi, “Deep learning based emotion analysis of microblog texts,” *Information Fusion*, vol. 64, pp. 1–11, 2020.

[22] T. Park, I. Ju, J. E. Ohs, and A. Hinsley, “Optimistic bias and preventive behavioral engagement in the context of covid-19,” *Research in Social and Administrative Pharmacy*, vol. 17, no. 1, pp. 1859–1866, 2021.
[23] T. W. Russell, J. T. Wu, S. Clifford, W. J. Edmunds, A. J. Kucharski, M. Jit et al., “Effect of internationally imported cases on internal spread of covid-19: a mathematical modelling study,” *The Lancet Public Health*, vol. 6, no. 1, pp. e12–e20, 2021.

[24] S. M. Fast, M. C. González, J. M. Wilson, and N. Markuzon, “Modelling the propagation of social response during a disease outbreak,” *Journal of The Royal Society Interface*, vol. 12, no. 104, p. 20141105, 2015.

[25] C. Cheng, “To be paranoid is the standard? panic responses to sars outbreak in the hong kong special administrative region,” *Asian Perspective*, pp. 67–98, 2004.

[26] D. Valdez, M. Ten Thij, K. Bathina, L. A. Rutter, and J. Bollen, “Social media insights into us mental health during the covid-19 pandemic: longitudinal analysis of twitter data,” *Journal of medical Internet research*, vol. 22, no. 12, p. e21418, 2020.

[27] M. O. Lwin, J. Lu, A. Sheldenkar, P. J. Schulz, W. Shin, R. Gupta, and Y. Yang, “Global sentiments surrounding the covid-19 pandemic on twitter: analysis of twitter trends,” *JMIR public health and surveillance*, vol. 6, no. 2, p. e19447, 2020.

[28] M. E. Basiri, S. Nemati, M. Abdar, S. Asadi, and U. R. Acharrya, “A novel fusion-based deep learning model for sentiment analysis of covid-19 tweets,” *Knowledge-Based Systems*, vol. 228, p. 107242, 2021.

[29] A. H. Sinclair, S. Hakimi, M. L. Stanley, R. A. Adcock, and G. R. Samanez-Larkin, “Pairing facts with imagined consequences improves pandemic-related risk perception,” *Proceedings of the National Academy of Sciences*, vol. 118, no. 32, 2021.

[30] M. Ali, N. Alam, and S. A. R. Rizvi, “Coronavirus (covid-19)—an epidemic or pandemic for financial markets,” *Journal of Behavioral and Experimental Finance*, vol. 27, p. 100341, 2020.

[31] M. Karl, F. Kock, B. W. Ritchie, and J. Gauss, “Affective forecasting and travel decision-making: An investigation in times of a pandemic,” *Annals of Tourism Research*, vol. 87, p. 103139, 2021.

[32] J. N. Druckman, S. Klar, Y. Krupnikov, M. Levendusky, and J. B. Ryan, “How affective polarization shapes americans’ political beliefs: A study of response to the covid-19 pandemic,” *Journal of Experimental Political Science*, vol. 8, no. 3, pp. 223–234, 2021.

[33] M. Roman and A.-P. Ploepeanu, “The effectiveness of the emergency elearning during covid-19 pandemic. the case of higher education in economics in romania,” *International Review of Economics Education*, vol. 37, p. 100218, 2021.

[34] K. Schmelz and S. Bowles, “Opposition to voluntary and mandated covid-19 vaccination as a dynamic process: Evidence and policy implications of changing beliefs,” *Proceedings of the National Academy of Sciences*, vol. 119, no. 13, p. e2118721119, 2022. [Online]. Available: https://www.pnas.org/doi/abs/10.1073/pnas.2118721119

[35] K. Barrafrem, G. Tinghög, and D. Västfjäll, “Trust in the government increases financial well-being and general well-being during covid-19,” *Journal of behavioral and experimental finance*, vol. 31, p. 100514, 2021.

[36] T. Christensen and P. Lægreid, “Balancing governance capacity and legitimacy: how the norwegian government handled the covid-19 crisis as a high performer,” *Public Administration Review*, vol. 80, no. 5, pp. 774–779, 2020.

[37] B. Steinegger, L. Arola-Fernández, C. Granell, J. Gómez-Gardeñes, and A. Arenas, “Behavioural response to heterogeneous severity of covid-19 explains temporal variation of cases among different age groups,” *Philosophical Transactions of the Royal Society A*, vol. 380, no. 2214, p. 20210119, 2021.

[38] G. W. Kołodko et al., “After the calamity: Economics and politics in the post-pandemic world,” *Polish Sociological Review*, vol. 210, no. 2, pp. 137–156, 2020.

[39] M. Leach, H. MacGregor, I. Scoones, and A. Wilkinson, “Post-pandemic transformations: How and why covid-19 requires us to rethink development,” *World Development*, vol. 138, p. 105233, 2021.

[40] A. C. McFarlane, “The aetiology of post-traumatic stress disorders following a natural disaster,” *The British Journal of Psychiatry*, vol. 152, no. 1, pp. 116–121, 1988.

[41] S. Goldrick-Rab, V. Coca, J. Gill, M. Peele, K. Clark, and E. Looker, “Self-reported covid-19 infection and implications for mental health and food insecurity among american college students,” *Proceedings of the National Academy of Sciences*, vol. 119, no. 7, p. e2111787119, 2022.

[42] E. Cambria, D. Das, S. Bandyopadhyay, and A. Feraco, “Affective computing and sentiment analysis,” in *A practical guide to sentiment analysis*. Springer, 2017, pp. 1–10.

[43] H. Kang, S. J. Yoo, and D. Han, “Senti-lexicon and improved naïve bayes algorithms for sentiment analysis of restaurant reviews,” *Expert Systems with Applications*, vol. 39, no. 5, pp. 6000–6010, 2012.
[44] V. Hatzivassiloglou and K. McKeown, “Predicting the semantic orientation of adjectives,” in 35th annual meeting of the association for computational linguistics and 8th conference of the association for computational linguistics, 1997, pp. 174–181.

[45] V. Hatzivassiloglou and J. Wiebe, “Effects of adjective orientation and gradability on sentence subjectivity,” in COLING 2000 Volume 1: The 18th International Conference on Computational Linguistics, 2000.

[46] W. Zhang, M. Wang, and Y.-c. Zhu, “Does government information release really matter in regulating contagion-evolution of negative emotion during public emergencies? from the perspective of cognitive big data analytics,” International Journal of Information Management, vol. 50, pp. 498–514, 2020.

[47] W. S. McCulloch and W. Pitts, “A logical calculus of the ideas immanent in nervous activity,” The bulletin of mathematical biophysics, vol. 5, no. 4, pp. 115–133, 1943.

[48] G. E. Hinton et al., “Learning distributed representations of concepts,” in Proceedings of the eighth annual conference of the cognitive science society, vol. 1. Amherst, MA, 1986, p. 12.

[49] Y. Li and T. Yang, “Word embedding for understanding natural language: a survey,” in Guide to big data applications. Springer, 2018, pp. 83–104.

[50] G. Liu and J. Guo, “Bidirectional lstm with attention mechanism and convolutional layer for text classification,” Neurocomputing, vol. 337, pp. 325–338, 2019.

[51] J. A. Hertz, Introduction to the theory of neural computation. CRC Press, 2018.

[52] R. Johnson and T. Zhang, “Supervised and semi-supervised text categorization using lstm for region embeddings,” in International Conference on Machine Learning. PMLR, 2016, pp. 526–534.

[53] P. Liu, X. Qiu, X. Chen, S. Wu, and X.-J. Huang, “Multi-timescale long short-term memory neural network for modelling sentences and documents,” in Proceedings of the 2015 conference on empirical methods in natural language processing, 2015, pp. 2326–2335.

[54] K. Jaseena and B. C. Kovoor, “Decomposition-based hybrid wind speed forecasting model using deep bidirectional lstm networks,” Energy Conversion and Management, vol. 234, p. 113944, 2021.

[55] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. Kaiser, and I. Polosukhin, “Attention is all you need,” in Advances in neural information processing systems, 2017, pp. 5998–6008.

[56] D. Bahdanau, K. Cho, and Y. Bengio, “Neural machine translation by jointly learning to align and translate,” arXiv preprint arXiv:1409.0473, 2014.

[57] X. Zhou, X. Wan, and J. Xiao, “Attention-based lstm network for cross-lingual sentiment classification,” in Proceedings of the 2016 conference on empirical methods in natural language processing, 2016, pp. 247–256.

[58] J. Du, L. Gui, R. Xu, and Y. He, “A convolutional attention model for text classification,” in National CCF conference on natural language processing and Chinese computing. Springer, 2017, pp. 183–195.

[59] G. Wang, C. Li, W. Wang, Y. Zhang, D. Shen, X. Zhang, R. Henao, and L. Carin, “Joint embedding of words and labels for text classification,” arXiv preprint arXiv:1805.04174, 2018.

[60] T. E. Trueman, E. Cambria et al., “A convolutional stacked bidirectional lstm with a multiplicative attention mechanism for aspect category and sentiment detection,” Cognitive Computation, vol. 13, no. 6, pp. 1423–1432, 2021.

[61] Y. Zhang and Q. Yang, “A survey on multi-task learning,” IEEE Transactions on Knowledge and Data Engineering, 2021.

[62] P. Liu, X. Qiu, and X. Huang, “Recurrent neural network for text classification with multi-task learning,” arXiv preprint arXiv:1605.05101, 2016.

[63] H. Linmei, T. Yang, C. Shi, H. Ji, and X. Li, “Heterogeneous graph attention networks for semi-supervised short text classification,” in Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP), 2019, pp. 4821–4830.

[64] H. Tang, D. Ji, C. Li, and Q. Zhou, “Dependency graph enhanced dual-transformer structure for aspect-based sentiment classification,” in Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics, 2020, pp. 6578–6588.

[65] K. Cho, B. Van Merriënboer, C. Gulcehre, D. Bahdanau, F. Bougares, H. Schwenk, and Y. Bengio, “Learning phrase representations using rnn encoder-decoder for statistical machine translation,” arXiv preprint arXiv:1406.1078, 2014.
[66] A. Graves, “Generating sequences with recurrent neural networks,” *arXiv preprint arXiv:1308.0850*, 2013.
[67] L. Zhang, Y. Lei, and Z. Wang, “Long-text sentiment analysis based on semantic graph,” in *2020 IEEE International Conference on Embedded Software and Systems (ICESS)*. IEEE, 2020, pp. 1–6.
[68] Y. Zhang and B. Wallace, “A sensitivity analysis of (and practitioners’ guide to) convolutional neural networks for sentence classification,” *arXiv preprint arXiv:1510.03820*, 2015.
[69] D. Asteriou, K. Pilbeam, and C. E. Pratiwi, “Public debt and economic growth: panel data evidence for asian countries,” *Journal of Economics and Finance*, vol. 45, no. 2, pp. 270–287, 2021.
[70] R. L. Greene, *The MMPI-2: An interpretive manual*. Allyn & Bacon, 2000.
[71] J. M. Wooldridge, “Inverse probability weighted m-estimators for sample selection, attrition, and stratification,” *Portuguese Economic Journal*, vol. 1, no. 2, pp. 117–139, 2002.
[72] H. White, “A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity,” *Econometrica: journal of the Econometric Society*, pp. 817–838, 1980.
[73] H. White and I. Domowitz, “Nonlinear regression with dependent observations,” *Econometrica: Journal of the Econometric Society*, pp. 143–161, 1984.
[74] S. D. Whitaker, “Did the covid-19 pandemic cause an urban exodus?” *Cfdd District Data Briefs*, no. cfdddb 20210205, 2021.
[75] E. E. Leamer, “Let’s take the con out of econometrics,” *The American Economic Review*, vol. 73, no. 1, pp. 31–43, 1983.
[76] B. Van Bockstaele, B. Verschuere, H. Tibboel, J. De Houwer, G. Crombez, and E. H. Koster, “A review of current evidence for the causal impact of attentional bias on fear and anxiety.” *Psychological bulletin*, vol. 140, no. 3, p. 682, 2014.
[77] J. Xu, G. Sun, W. Cao, W. Fan, Z. Pan, Z. Yao, and H. Li, “Stigma, discrimination, and hate crimes in chinese-speaking world amid covid-19 pandemic,” *Asian journal of criminology*, vol. 16, no. 1, pp. 51–74, 2021.
[78] S. Choi, J. Y. Hong, Y. J. Kim, and H. Park, “Predicting psychological distress amid the covid-19 pandemic by machine learning: discrimination and coping mechanisms of korean immigrants in the us,” *International journal of environmental research and public health*, vol. 17, no. 17, p. 6057, 2020.
[79] P. Slovic, “Perception of risk,” *Science*, vol. 236, no. 4799, pp. 280–285, 1987.
[80] Y. Ji, M. Qi, and W. Qi, “The effect path of public acceptance and its influencing factors on public willingness to participate in nuclear emergency governance,” *International Journal of Disaster Risk Reduction*, vol. 71, p. 102806, 2022.