Data Mining Project: Impact of the COVID-19 on GDP of various industries in Canada

Carson Kai-Sang Leung
Department of Computer Science
University of Manitoba
Winnipeg, MB, Canada
kleung@cs.umanitoba.ca

Yipeng Liu
Department of Computer Science
University of Manitoba
Winnipeg, MB, Canada
liuy3468@myumanitoba.ca

***Abstract*—Gross Domestic Product (GDP) is used to estimate a country's growth rate and its economy's size. Due to the global outbreak of COVID-19, many countries have put measures to slow its spread. Measures to mitigate the spread of the COVID-19 include company layoffs and business closures, which impacted GDP. Since business activities that generate income have been closed or slowed down, Canada becomes one of the countries that have been affected by the outbreak of COVID-19. In this paper, we use both linear regression and polynomial regression models to predict and analyze the impact of the COVID-19 on the GDP of various industries in Canada, which helps users better understand the impact of the COVID-19 on Canada. Evaluation results show that the impact of the COVID-19 on GDP is different among Canadian industries. The GDP of the Canadian accommodation and food services industry, Canadian arts, entertainment and recreation services industry, and Canadian manufacturing industry is significantly affected by COVID-19. However, the GDP of Canadian agriculture, forestry, fishing and hunting industries, Canadian utilities, and the Canadian cannabis sector are almost unaffected by COVID-19.**

Keywords—Analysis, Canadian industry, COVID-19, data mining, Gross Domestic Product (GDP), prediction, regression

I. INTRODUCTION

The spread of COVID-19 has forced a sharp decline in global economic activity. As one of the largest economies, Canada's economy has also suffered an unprecedented hit. Coronavirus, which causes COVID-19, can spread through close contact from an infected person to others, including droplets, airborne, surface or fecal-oral transmission [1]. The high infectivity and morbidity of the virus forced the government to take quarantine measures and containment measures to quickly contain the source of infection of the virus and stop its spread [2]. However, the wide range of prolonged lockdown, quarantine, and enforced stay-at-home orders caused many businesses and individuals to spend less, triggering a recession [3]. The lack of orders and income of enterprises led to large-scale layoffs and increased unemployment.

The percentage change in the gross domestic product (GDP) of a year is an essential and closely-watched indicator of a country's economic performance [4]. A lower GDP can reflect that there is less wealth and fewer job opportunities in this country. Data from *Statistics Canada* confirm that the recession began at the same time as the national lockdown. Closures of businesses, schools and borders in the last two weeks of March contributed to a 7.2% drop in GDP for the whole month [5]. Almost no industry is immune to this economic downturn.

In this paper, we use regression models to analyze the impact of the COVID-19 on the GDP of various industries in Canada. First, we analyze the trend of Canada GDP changes from January 1997 to December 2020. The result shows that Canada's GDP has been on a roughly upward trend from January 1997 to December 2020. Then, we focus on the COVID-19 pandemic period. We analyze the impact of the COVID-19 on the GDP of various industries in Canada. By evaluating the result, we find that: (a) both of the production and service industry were affected by COVID-19, (b) the sectors that have been affected by the pandemic can be divided into two groups: severely affected sectors and the sectors that affected to a lesser extent.

In the production industry, the manufacturing industry was the industry that suffered the most in the first few weeks after the crisis broke out [6]. According to the monthly survey of the manufacturing industry, nearly 80% of companies in the industry reported a slowdown in their activities [6]. In particular, the automotive industry has been hit hard by factory closures. In the service sector, food, accommodation and the arts, entertainment and leisure sectors have been most affected. The jobless rate rose 2.2 percentage points to 7.8% in March, which is the most significant one-month increase on record [7]. During this time, economic uncertainty and loss of income limited consumers' willingness to spend their money, which led to further declines in GDP for a while in the future.

However, there are a few industries that almost unaffected by the recession. The utility industry only had a slight decline in April, which is 1.8% [8]. There is no significant shift in overall power consumption since the residential load demand increased due to the stay-at-home orders. Also, the economic downturn had no apparent impact on the agriculture sector. Due to the closure of restaurant and bar, more consumers choose to cook at home instead of eating out, which contribute to an increase in sale volume of agriculture products at grocery stores. Surprisingly, there is a 17% growth in GDP in the cannabis sector because citizens spend more time at home instead of travelling during the lockdown [9].

Our *key contributions* of this paper include our construct models for analyzing the impact of the COVID-19 on the GDP of various industries in Canada. Our models include:

- Linear Regression Model for analyzing the trend of Canada GDP changes from January 1997 to December 2020.
- Polynomial Regression Model for analyzing the trend of Canada GDP changes from January 1997 to December 2020.
- Linear Regression Model for analyzing the impact of the COVID-19 on GDP of specific Canadian sector.

- Polynomial Regression Model for analyzing the impact of the COVID-19 on GDP of specific Canadian sector.

Our models and the evaluation can help users better understand the impact of the COVID-19 on Canada. The evaluation result illustrates to users how Canadian sectors were affected during the outbreak and explained why. Also, the result gives users some insight into the trends in the job market in Canada during the COVID-19 period and can guide people looking for work during the COVID-19 period. It may even be helpful to the Canadian government in macro-control of various industries during the COVID-19 period.

We organize the remainder of this paper as follows. The following section provides a briefly background information about the impact of COVID-19 on GDP and analyzes some related works. Section III explains the algorithms we use in this paper. Section IV shows our evaluation results. In the end, sections V shows our conclusion and future work.

II. Background and Related Works

As the coronavirus spread globally, every aspect of people's lives has been affected, including social lifestyle, physical and mental health, working environment, the personal financial situation. Some researchers have conducted some work on COVID-19:

- In order to better understand and come up with effective countermeasures, there have been researched papers about prediction [10], [11] and analysis of COVID-19 outbreak [12], [13].
- To explore more effective tools to predict the disease and help doctors to make decisions, there have been studies about predicting COVID-19 based on chest x-rays [14], [15], [16], [17].
- For a better result of preventing the spread of coronavirus, there have been studies works on analyzing the effectiveness of different ways of prevention [18], [19], [20].

COVID-19 pandemic is not only threatening people's health condition but also ruining their financial conditions. Due to the outbreak of this epidemic, a series of isolation and blockade measures taken by governments at all levels have exerted a severe impact on business operation, individual employment, family consumption and other economic aspects. Many scholars have researched the economic impact of COVID-19 as well.

Since the government needs to make informed policies to exit from the economic crisis brought by the COVID-19 pandemic, analyzing the economic position during the COVID-19 period is essential. For instance, Iacus et al. focus on studying the impact of the air travel ban on the world economy under the influence of the COVID-19 [21]. They collect and prepare data on air passenger flow worldwide to analyze the losses of travel restriction on the aviation sector. Their prediction model uses a non-homogeneous Poisson process with a periodic intensity function calibrated on historical data and predicts the future data in a non-linear way. Unlike Iacus et al.'s work on analyzing and predicting the impact of the air travel ban on world GDP, we focus on analyzing and predicting the impact of the COVID-19 on the GDP of various industries in Canada. We are committed to analyzing how the GDP changes among different Canadian industries during the COVID-19 period, which will help users better understand the impact of the COVID-19 on Canada. Even help the government adopt different macro-control measures for different industrial chains.

Among related works on analyzing the GDP data during the COVID-19 pandemic period, many focused on building more accurate prediction models. For instance, Jena et al. use a multilayer artificial neural network model to predict the GDP values for the second quarter of 2020 for the United States, Mexico, Germany, Italy, Spain, France, India, and Japan [22]. They achieved a prediction error of less than 2% by the model during the testing procedure. Unlike Jena et al. work on building a more accurate prediction model, we focus on using simpler regression models to predict the GDP value, then analyze the result. Based on the prediction results, we conclude that the GDP of most Canadian industries has a decreasing trend during the COVID-19 period. However, some Canadian industries are almost unaffected by COVID-19. We also analyze the reason for the GDP trend of these industries.

III. Algorithms

We first build three linear regression models to evaluate our topic. Then, we build three polynomial regression models to see if our linear regression models are accurate enough to predict our topic.

In Section III-*A* and Section III-*B*, we will introduce the construction methods of the two regression models separately. In Section III-*C*, we will introduce the instances of our models. In Section III-*D,* we will introduce the metric that measure the accuracy of our models.

A. Linear Regression Algorithm

Given an example with d attribute descriptions $x = (x1, x2,..., xd)$, where i is the value of the *ith* attribute. The linear regression model tries to learn a function by a linear combination of attributes then use it to make a prediction:

$$f(x) = w_1x_1 + w_2x_2 + ... + w_dx_d + b$$

Since we only focus on one attribute, we will use the unary linear regression model for our project. Its formula is:

$$\hat{y} = wx_i + b$$

Each terms in this equation are:

- $\hat{y}$ denotes the predicted response value for the experimental x.
- x_i denotes the predictor value (independent variable) given by our data file.
- w denotes the regression coefficients of this model.
- b denotes the error term of this model.

To make our model more accurate, we need to find a proper value of w to minimize the value b. In our project, we use the least square method to find the value of w and b:

$$E(w, b) = \sum_{i=1}^{n} (wx_i + b - y_i)^2$$

Take the partial derivative of the above equation and set them equal to zero, we can get:

$$w = \frac{\sum_{i=1}^{n} y_i(x_i - \bar{x})}{\sum_{i=1}^{n} x_i^2 - \frac{1}{n}\left(\sum_{i=1}^{n} x_i\right)^2}$$

$$b = \frac{1}{n}\sum_{i=1}^{n}(y_i - wx_i)$$

B. Polynomial Regression Algorithm

Polynomial regression is a kind of linear regression because the regression function is linear concerning the regression coefficient. The most significant advantage of polynomial regression is that it can deal with many nonlinear problems. It plays an essential role in regression analysis because piecewise polynomials can approximate any function. In our project, we use unary polynomial regression. Its formula is:

$$\hat{y} = b_0 + b_1 x + b_2 x^2 + \ldots + b_m x^{\mathrm{m}}$$

Using polynomial regression, we added new features to the data set, which are the polynomial combinations of the original features. We also use the same technique as we used in linear regression to find the proper value of w and b to make our model more accurate.

C. Instances of our models

Our project contains three model sets, and each model set contains two models (one linear regression model and one polynomial regression model) for analyzing the same subtopic.

- The first model set is used to analyze and predict the trend of Canadian GDP from January 1997 to December 2020. We choose the monthly timeline as the independent variable and choose the Canadian GDP value as the dependent variable in collecting data. For building the models, we first split the data set into training data and test data in an 8:2 ratio, and we will not training the data of the COVID-19 period.
- The second model set is similar to the first model set, except we will train the data of the COVID-19 period.
- The third model set is used to analyze and predict the impact of the COVID-19 on the GDP of various industries in Canada. In collecting data, we choose the COVID-19 new cases rate as the independent variable and choose the changes of GDP value of various Canadian industries as the dependent variable. For building the models, we split the data set into training data and test data in a 5:5 ratio because there only have a few data during the COVID-19 period.

D. R-squared value

In this paper, we will use R-squared value as the metric to measure the accuracy of our models. R-squared value is the coefficient of determination of a regression model. It is a statistical measure to reflect the goodness of fit of the model. The value of R-squared normally ranges from 0 to 1. The larger R-squared value is (close to 1), the better the regression equation is. Suppose there is a Baseline Model simply predict that all the $\hat{y}_i$ values are equal to the mean of the dependent variable, which is not accurate. If the the R-squared value of our model equals to 1, it means all the predicted values are equal to the true values. If the the R-squared value of our model equals to 0, it means our model is equals to the Baseline Model. If the the R-squared value of our model less than 0, it means our model is worse than the Baseline Model. The equation of R-squared value is:

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(\hat{y} - y_i)^2}{\sum_{i=1}^{n}\left(\bar{y} - y_i\right)^2}$$

- The numerator represents the prediction errors of our model.
- The denominator represents the prediction errors of the Baseline Model.

IV. Evaluation

This section will use our prediction models to analyze the impact of the COVID-19 on Canada's GDP. In section IV-*A*, we analyze Canada's GDP changes from January 1997 to December 2020, which contains ten months of COVID-19 period from March 2020 to December 2020. In section IV-*B*, we analyze the relationship between the COVID-19 new cases rate and Canadian industry GDP value changes. All of our data are monthly data.

A. Trend of Canada GDP changes

We use two model sets to analyze the trend of Canada's GDP changes. The training data of the first model does not include the data of the COVID-19 period. In contrast, the training data of the second model includes the data of the COVID-19 period. In both of these two models, we use the monthly timeline from January 1997 to December 2020 as the independent variable and use Canada's GDP value during this period as the dependent variable. We collect the data from *Statistic Canada* [23].

1) Model with training data from COVID-19 period not included

We first build a linear regression model to analyze the trend of Canada's GDP changes. When building this model, we only read the data of the non-COVID-19 period from the data file. We then split the data into a training set and a test set with random selection in an 8:2 ratio. Now, the training set does not contain the data of the COVID-19 period. Then, we train the model. Before we use this model to make the prediction, we add the COVID-19 period to the test set to get the complete prediction curve from January 1997 to December 2020. The trend of Canada GDP changes (training data only from the non-COVID-19 period) sees Fig. 1.

Fig. 1 shows Canada's GDP has been on a roughly upward trend from January 1997 to December 2020. In October 2008, Canada was hit by the financial crisis [24], and its GDP declined for seven months, ended up down 4.52%. The coronavirus disease hit Canada's GDP harder than the financial crisis, knocking 17.75% off it, but the decrease of Canada's GDP only keep for two months. Overall, the crisis brought by the COVID-19 pandemic is the biggest one in Canadian history that has affected Canadian GDP. Measured by the GDP decline alone, the decline brought from COVID-19 was more than four times as big as the financial crisis. However, Canada's GDP recovered faster at this time, which took only two months.

We use the R-squared value to evaluate the accuracy of this model. Since the R-squared value of the left graph in Fig. 1 is larger than 0.8, we conclude our model performs well on predicting the Canadian GDP trend from January 1997 to

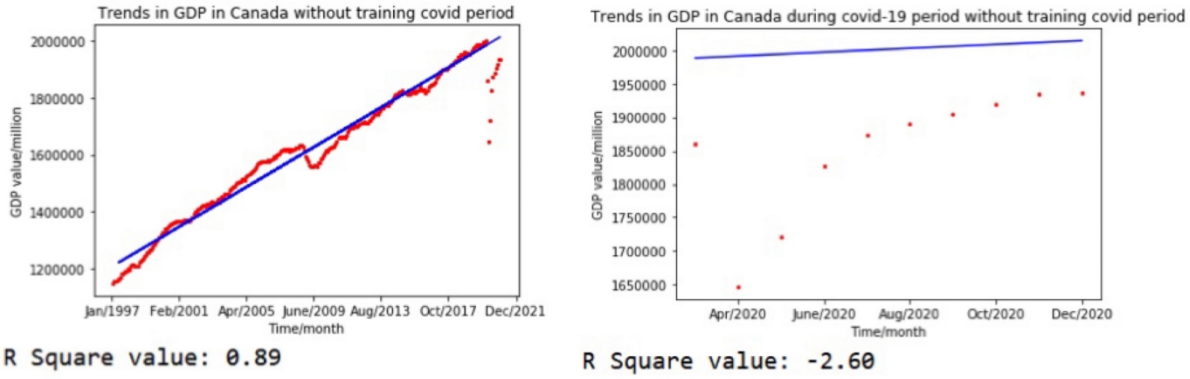

Fig. 1. Trends in GDP of Canada from Jan.1997 to Dec.2020 without training the data of COVID-19 period (left) & Trends in GDP of Canada during COVID-19 period without training COVID-19 data (right), both using linear regression model.

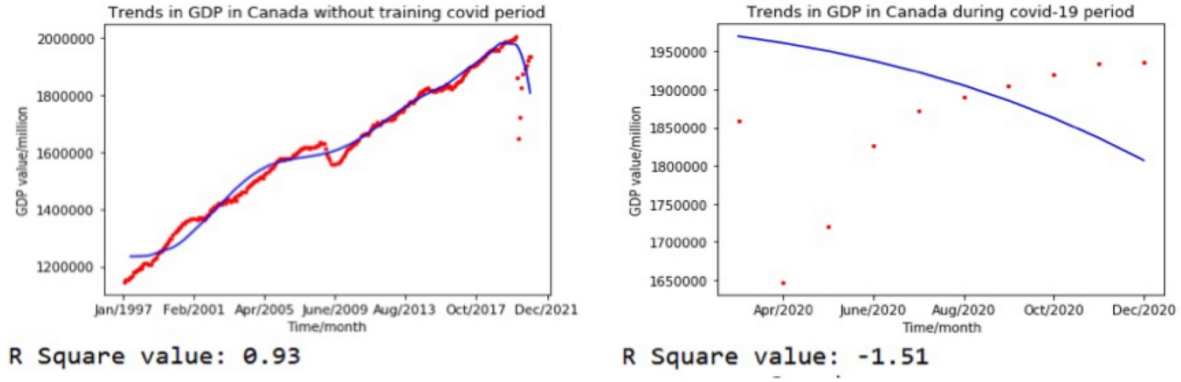

Fig. 2. Trends in GDP of Canada from Jan.1997 to Dec.2020 without training the data of COVID-19 period (left) & Trends in GDP of Canada during COVID-19 period without training COVID-19 data (right), both using polynomial regression model.

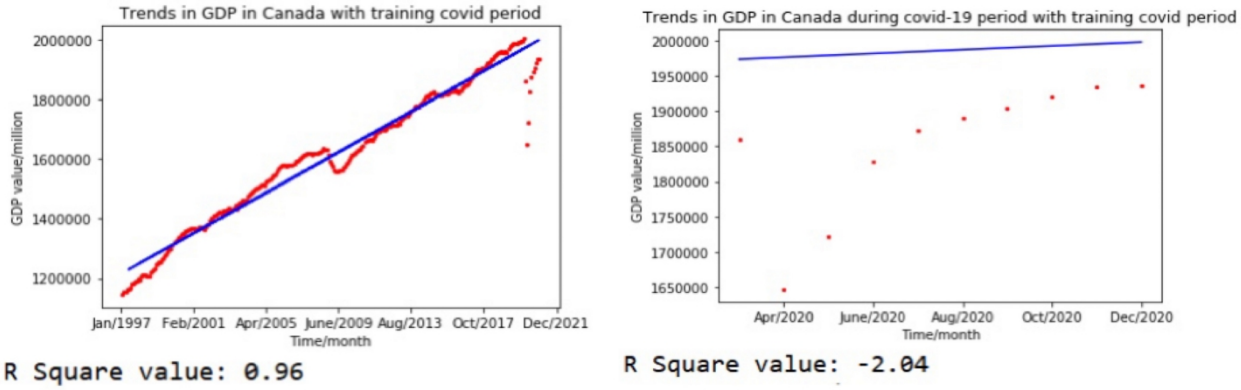

Fig. 3. Trends in GDP of Canada from Jan.1997 to Dec.2020 with training the data of COVID-19 period (left) & Trends in GDP of Canada during COVID-19 period with training the data of COVID-19 period (right), both using linear regression model.

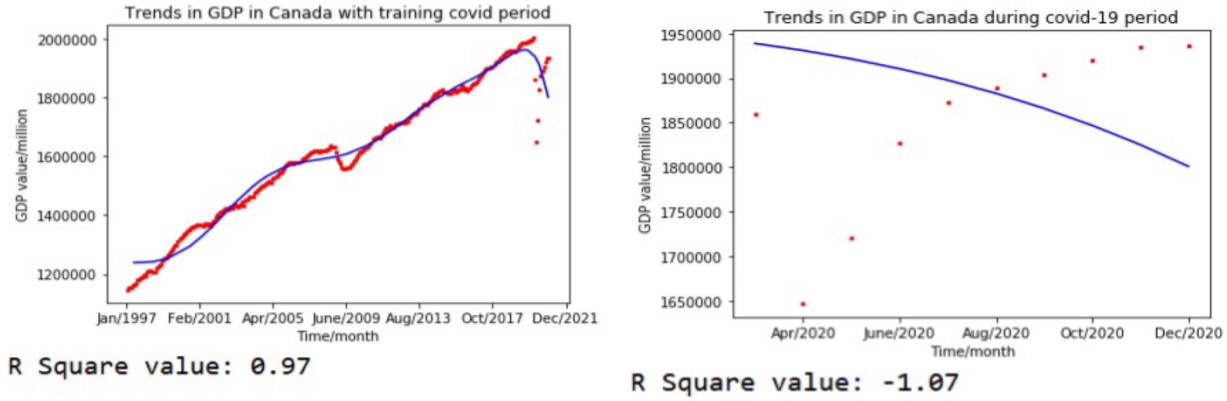

Fig. 4. Trends in GDP of Canada from Jan.1997 to Dec.2020 with training the data of COVID-19 period (left) & Trends in GDP of Canada during COVID-19 period with training the data of COVID-19 period (right), both using polynomial regression model.

December 2020. It means that our prediction results are accurate, which has a high degree of credibility. However, the R-squared value of the right graph in Figure 1 is -2.60, which shows that our prediction line does not correlate with the data of the COVID-19 period. This result is predictable because we did not put the data of the COVID-19 period into the training set. So our prediction line performs poorly on the data during the COVID-19 period.

We also build a polynomial model to analyze the trend of Canada's GDP changes. We used the same technique used in the previous linear regression model to read and split the data set about this model. Then we sort both the training set and test set in ascending order. It helps our model to produce a coherent graph. For doing polynomial regression, we need to make the independent variable polynomial. After many attempts, we find that the best degree of the polynomial is 9. It means our polynomial regression model performs the best when the degree of the polynomial is 9. The result of our prediction sees Fig. 2.

Since the R-squared value of both graphs in Fig. 2 is larger than the R-squared value of the corresponding graph in Fig. 1, we conclude that the polynomial regression model performs better than the linear regression model on predicting the trend in Canadian GDP.

2) Model with training data from COVID-19 period included

In this model set, we will put the data of COVID-19 period into training data. With training the data of the COVID-19 period, we will see if our predictions are more accurate than previous models.

We first build a linear regression model to analyze the trend of Canada's GDP changes. When building this model, we read the complete data from the data file. Then, we split the data into a training set and a test set with random selection in an 8:2 ratio. Then, we train the model. The result of our prediction sees Fig. 3.

Compare to Fig. 1, the R-squared value of both the two graphs in Fig. 3 is larger than the R-squared value of the corresponding graph in Fig. 1. It shows that with training the data of the COVID-19 period, the linear regression model performs better than not training the COVID-19 data on predicting the trends in GDP of Canada.

We also build a polynomial model to analyze the trend of Canada's GDP changes. We used the same technique to process the data set about this model as we did to build the previous linear regression model. Then we sort both the training set and test set in ascending order. It helps our model to produce a coherent graph. After many attempts, we find the best degree of the polynomial is 9. The result of our prediction sees Fig. 4.

Since the R-squared value of both graphs in Fig. 4 is larger than the R-squared value of the corresponding graph in Fig. 3, we conclude that the polynomial regression model performs better than the linear regression model on predicting the trend in Canada's GDP. Also, since the R-squared value of both graphs in Fig. 4 is larger than the R-squared value of the corresponding graph Fig. 2, we conclude that training the data of the COVID-19 period can make the polynomial regression model performs better on predicting the trends in Canada's GDP.

This model is the most accurate one in this section, but the R-squared value of the right graph in Fig. 4 is still negative. It means our prediction line does not correlate with the data of the COVID-19 period. We summarize the reasons why our models have an inaccurate prediction of Canadian GDP value during the epidemic:

- Because of the impact of the COVID-19 disease, Canada's GDP during the COVID-19 period is very different from the previous trend.
- Also, the data of the COVID-19 period is only a tiny part of our data set, so our model treats these data as bias when training the data set.

In the next section, we will focus on the data of the COVID-19 period, and those models will have a more accurate prediction of the changes of GDP value during the COVID-19 period.

B. Impact of the COVID-19 on GDP of various industries in Canada

In this section, we build a model set to analyze the impact of the COVID-19 on the GDP of various industries in Canada. By evaluating the prediction results, we divide the industries into two groups. The first group is the industries whose GDP is significantly affected by the COVID-19. The GDP of the severely affected industries fell by around 60%. The second group is the industries that are affected to a lesser extent. One of them even had a growing trend on GDP. In each model, we use the COVID-19 new cases rate as the independent variable and use the changes of GDP value of one specific Canadian industry as the dependent variable.

1) Data Collection and Integration

For building our models, we need to find data to make our independent variable and dependent variable:

- We gathered the monthly GDP value in the COVID-19 period of various Canadian industry from *Statistic Canada* [23]. We choose the monthly GDP value in COVID-19 period of a specific industry as our dependent variable.
- We gathered the COVID-19 data from *Our World in Data* [25]. We choose the data of the COVID-19 new cases of Canada from the COVID-19 data. Since the GDP values are monthly data, we make the new COVID-19 cases in each month as our independent variable.
- We combine the data of these two variables in a data file.

2) Data Preprocessing

To verify that the selected variables are correlated, we need to calculate the correlation coefficient between the variables. We choose the data file of the Canadian goods-producing industry for testing the correlation. According to preliminary observation, the GDP of the Canadian goods-producing industry decreased significantly under the influence of the COVID-19, so the variables should have a strong negative correlation. However, the correlation coefficient of these two variables is 0.174, which shows a very poor positive correlation between our variables. It is the opposite of what we expected. We decide to preprocess the data. We find the COVID-19 new cases rate and the GDP value changes are reasonable variables after some attempt.

We calculate the COVID-19 new cases rate by using the total new cases of each month divide by the total new cases of last month. We treat the results as the independent variable. We calculate the GDP value changes by using the GDP value of each month minus the GDP value of last month. We treat the result as the dependent variable. By computing, the correlation coefficient of these two variables is -0.919, which shows a strong negative correlation between the variables, as we desired.

Then, we drop the bias records, which position far from the other spots, and the empty records for making our model more accurate. We also sort the data file in ascending order before we build the model. It helps our model to produce a coherent graph.

3) Significantly Affected Industries

This part will analyze the GDP changes of Canadian industries whose GDP are significantly affected by COVID-19 pandemic. We selected the three most representative industries. We will show how COVID-19 affects the GDP of Canadian accommodation and food services industries, the Canadian arts, entertainment and recreation services industry, and the Canadian manufacturing industry.

Firstly, we analyze the difference in the trend of the GDP of the industry during the epidemic and non-epidemic periods. The result shows in Fig. 5.

From Fig.5, we can see that the trend of the GDP of these three industries during COVID-19 period are completely different from their GDP trend in non-COVID-19 period. This shows the GDP of these industries are significantly affected by COVID-19 pandemic.

Secondly, we use a linear regression model and a polynomial regression model to analyze how COVID-19 affects the GDP of Canadian accommodation and food services industries, the Canadian arts, entertainment and recreation services industry, and the Canadian manufacturing industry.. Recall from early, we use the COVID-19 new cases rate as the independent variable and use the GDP value changes of a specific Canadian industry as the dependent variable. Since our data set is small, we split the data set into a training set and a test set with a 5:5 ratio in both linear regression and polynomial regression models. For the linear regression model, we can then just training the model and make the prediction. However, we need to sort both the training set and test set in ascending order before making our prediction. It helps our model to produce a coherent graph.

The result of our prediction by linear regression model see Fig. 6.

By observing Fig. 6, we can see our prediction model performs well with the Canadian industries, in which GDP is significantly affected by COVID-19. It shows that the COVID-19 pandemic has indeed had an impact on these industries, causing their GDP to fall seriously. We analyze the reasons why the GDP value of these industries decrease hard during the COVID-19 period.

1. Canadian accommodation and food services industries

- As one of the severely affected sectors, accommodation services suffered particularly the most, which plunged 45.7% in April 2020, owing to the restriction on international and provincial travel [8]. Food service shrank by 40.8% due to a wide range of prolonged lockdown, quarantine and enforced stay-at-home orders [8].
- Many of these food services businesses closed temporarily or permanently or operated at a significantly reduced capacity [26]. Also, as customers replaced eating at the restaurant with ordering online, take-out and delivery services become most restaurants and bars' sole source of income.

2. Canadian arts, entertainment and recreation services industries

- Due to the restriction rules by provincial governments, arts, entertainment and recreation, have also experienced significant economic setbacks, which fell 25.6% in April 2020. Economically, arts and culture contribute 3.8% of Canada's GDP [27]. However, during the COVID-19 pandemic, many of the protective actions implemented in adherence to the Canadian government for Emergency Management and Civil Protection Act (EMCPA) on social distancing forced the closure of non-essential businesses and facilities that directly support artistic and athletic endeavours [28].

3. Canadian manufacturing industry

- Due to the same reason, some of the non-durable manufacturing temporarily closed their door, which leads to a 15.4% decrease at the same time [8].
- Durable manufacturing fell 29.2%, in which motor vehicle was one of the sectors worst hit by the disasters (-97.7%) [8]. Since there is a great interdependence between the Canadian and U.S. auto industries [29], the continued closure of automotive plants on both sides of the Canada-U.S. border has hit the auto industry hard. Also, the 75% flight suspension rate of international airlines and lower demand for domestic flight lead to an aerospace product, and parts manufacturing was down by 10.6% [30].

Then, we construct a polynomial regression model to analyze how COVID-19 affects the GDP of these three industries, whose GDP is significantly affected by COVID-19. We choose the polynomial degree as two because when we increase the degree of the polynomial higher than 2, our model will get overfitting. When the model is overfitting, although the prediction result of the training set is getting better, the prediction result of the test set will worsen. The result of our prediction by polynomial regression model see Fig. 7.

We can see that the accuracy of the polynomial regression model and the linear regression are pretty similar. It is because our data set has an almost linear relationship.

4) Hardly Affected Industries

This part will analyze the GDP changes of Canadian industries whose GDP are hardly affected by COVID-19 pandemic. We selected the three most representative industries. We will show how COVID-19 affects the GDP of Canadian agriculture, forestry, fishing and hunting industries, Canadian utilities, and the Canadian cannabis sector.

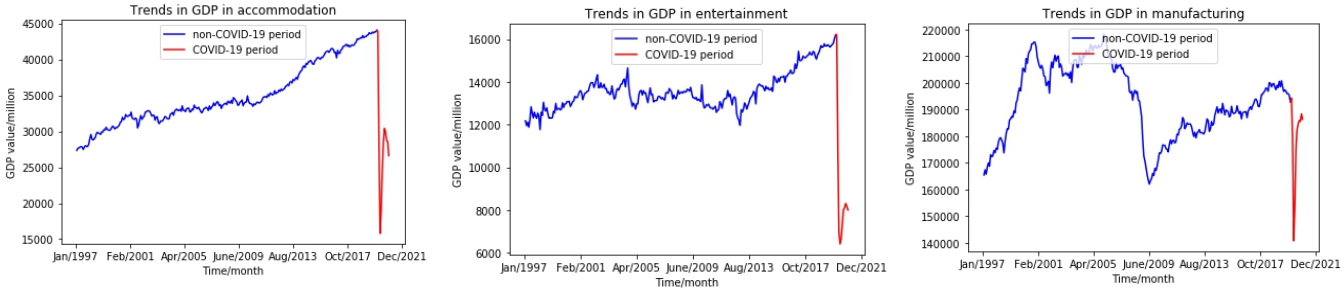

Fig. 5. Trends in GDP of three significantly affected industries from Jan.1997 to Dec.2020. Canadian accommodation and food services industries (left), Canadian arts, entertainment and recreation services industry (middle), and Canadian manufacturing industry (right).

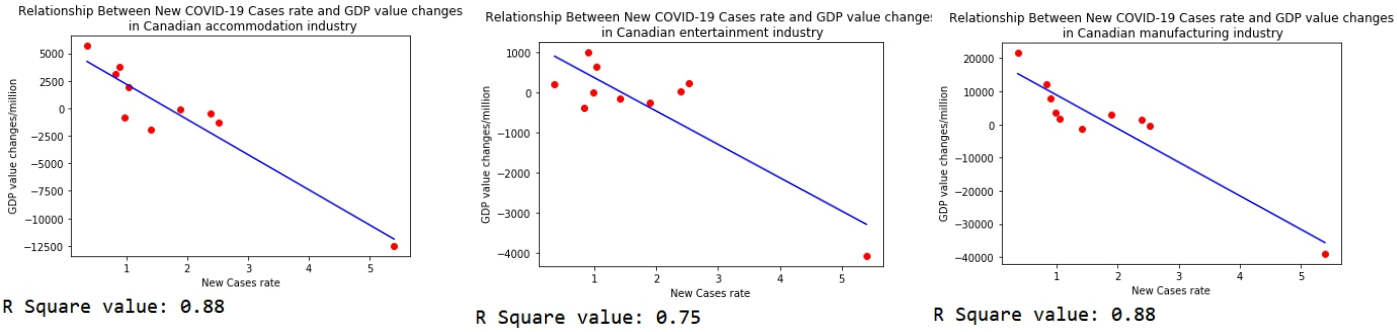

Fig. 6. The relation ship between new COVID-19 cases and GDP value changes of three significantly affected industries. Canadian accommodation and food services (left), Canadian arts, entertainment and recreation industries (middle), Canadian manufacturing (right), both using linear regression model)

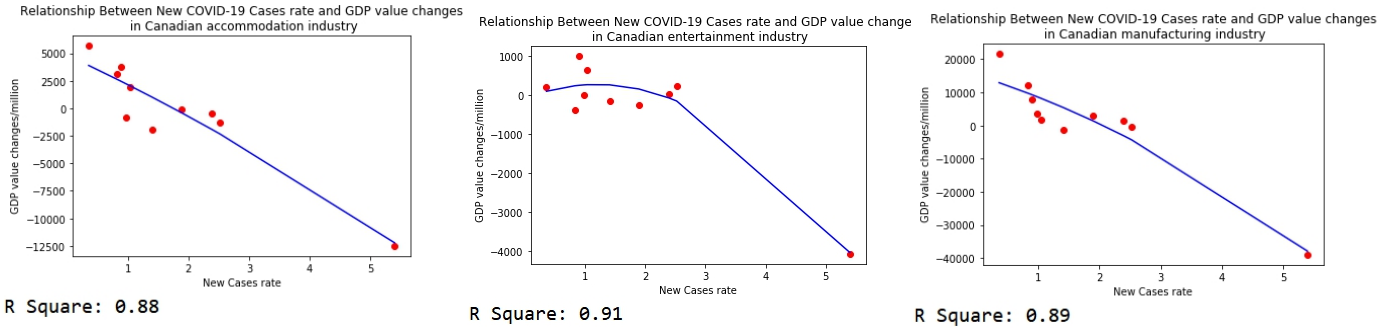

Fig. 7. The relation ship between new COVID-19 cases and GDP value changes of three significantly affected industries. Canadian accommodation and food services (left), Canadian arts, entertainment and recreation industries (middle), Canadian manufacturing (right), both using polynomial regression model)

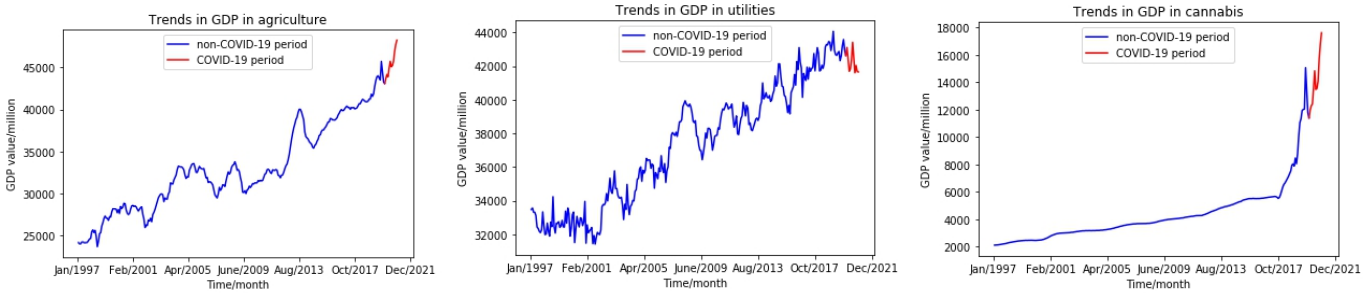

Fig. 8. Trends in GDP of three hardly affected industries from Jan.1997 to Dec.2020. Canadian agriculture, forestry, fishing and hunting industries (left), Canadian utilities (middle), and the Canadian cannabis sector (right).

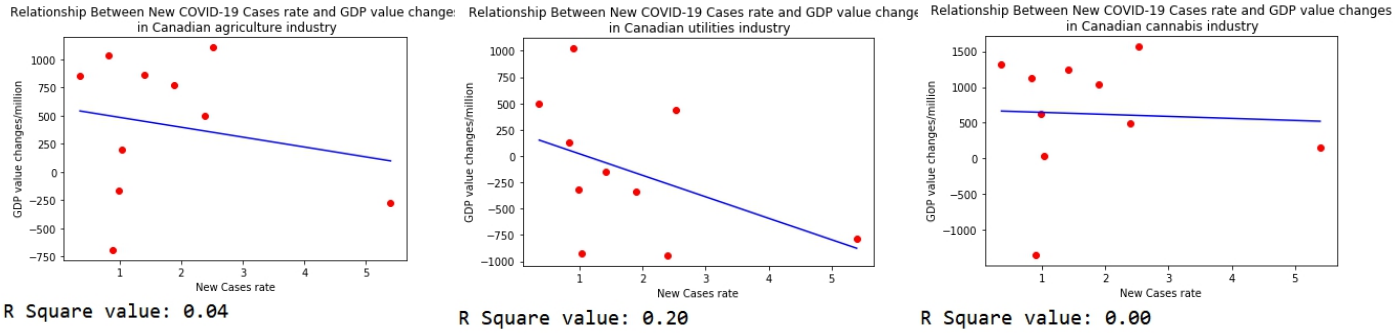

Fig. 9. The relation ship between new COVID-19 cases and GDP value changes of three hardly affected industries.Canadian agriculture, forestry, fishing and hunting industries (left), Canadian utilities (middle), Canadian cannabis sector (right), both using linear regression model)

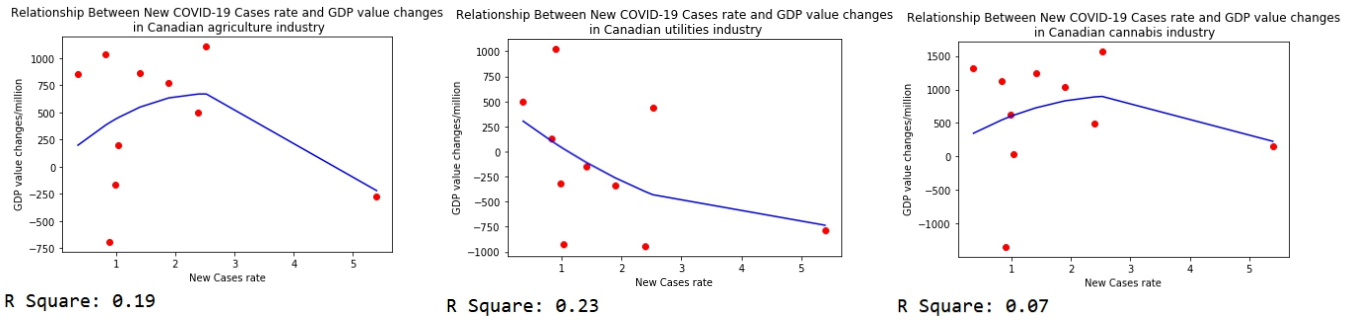

Fig. 10. The relation ship between new COVID-19 cases and GDP value changes of three hardly affected industries. Canadian agriculture, forestry, fishing and hunting industries (left), Canadian utilities (middle), Canadian cannabis sector (right), both using polynomial regression model)

Firstly, we analyze the difference in the trend of the GDP of the industry during the epidemic and non-epidemic periods. The result shows in Fig. 8.

From Fig. 8, we can see that the trend of the GDP of these three industries during COVID-19 period are almost as same as their GDP trend in non-COVID-19 period. Canada cannabis sector, in particular, has not seen a decline in GDP and has maintained its previous upward trend. This shows the GDP of these industries are almost unaffected by COVID-19 pandemic.

We use the same models as we used in section IV-*B-3)* to analyze how COVID-19 affects the GDP of Canadian agriculture, forestry, fishing and hunting industries, Canadian utilities, and the Canadian cannabis sector.

The result of our prediction by linear regression model see Fig. 9.

By observing Fig. 9, we can see that the R-squared value of our prediction model is small. It shows our prediction model performs poorly with the Canadian industries whose GDP is almost unaffected by COVID-19. Because the COVID-19 disease affects a minor influence on these industries, the relationship between our independent variable and the dependent variable is weak. We analyze the reasons why the GDP value of these industries almost unaffected during the COVID-19 period.

1. Canadian Agriculture, forestry, fishing and hunting industries

- As a significant contributor to Canada's economy, the agriculture sector has held up well in 2020 [31]. As customers replaced eating out with cooking at home, grocery store revenue had a significant increase. The agriculture sector had one of the smallest declines in February and then surged since the lockdown started in March [32].

2. Canadian Utilities

- As a provider of critical infrastructure, power and utility industries were prepared to respond to and plan for emergencies, including health emergencies. The COVID-19 pandemic leads to a major lifestyle-changing globally, most Canadian companies allowing employees to work from home full time. Therefore, even though there was a significant decrease in industrial load demand, the residential load demand vastly increased [33].

3. Canadian Cannabis sector

- Unlike most other sectors reporting negative growth, the cannabis sector saw almost 17% growth in GDP through 2020 [9], mainly because consumers spend more time at home than travelling outside during the lockdown.

In our polynomial regression model, we still choose the polynomial degree as two to avoid overfitting. The result of our prediction by polynomial regression model see Fig. 10.

We can see that even the accuracy of our polynomial model is higher than the linear regression model, it still very low. It demonstrates again that the COVID-19 has little impact on the GDP of these industries.

V. Conclusion and Future Work

In this paper, we present regression models to analyze the impact of the COVID-19 on GDP of various industries in Canada. Our evaluation result shows that the Canadian sectors that have been affected by the pandemic can be divided into two groups: severely affected sectors and the sectors that affected to a lesser extent. We selected the three most representative industries from each group for analysis and explained why their GDP changed. As future work, we plan to build more accurate models to make a more accurate prediction on the GDP value of each Canadian sector.

References

[1] Statistics Canada. (2021, Mar. 21). COVID-19: Main modes of transmission [Online]. Available: https://www.canada.ca/en/public-health/services/diseases/2019-novel-coronavirus- infection/health-professionals/main-modes-transmission.html

[2] L. Jones, D. Palumbo, D. Brown. (2021, Jan. 21). Coronavirus: How the pandemic has changed the world economy [Online]. Available: https://www.bbc.com/news/business-51706225

[3] C. Baek, P. B. McCrory, T. Messer, and P. Mui. (2020, July). "Unemployment Effects of Stay-at-Home Orders: Evidence from High Frequency Claims Data". IRLE Working Paper No. 101-20. Available: http://irle.berkeley.edu/files/2020/07/Unemployment-Effects-of-Stay-at- Home-Orders.pdf

[4] T. Callen. (2008, Dec). "Back to Basics: What Is Gross Domestic Product? ". Finance & Development [Online]. Available: http://purochioe.rrojasdatabank.info/imfongdp.pdf

[5] Statistics Canada. (2020, May. 29). Gross domestic product by industry, MARCH 2020 [Online]. Available: https:// www150.statcan.gc.ca/n1/daily-quotidien/200529/dq200529b-eng.htm

[6] Statistics Canada. (2020, May. 14). Monthly survey of Manufacturing, MARCH 2020 [Online]. Available: https://www150.statcan.gc.ca/n1/daily-quotidien/200514/dq200514a-eng.htm

[7] Statistics Canada. (2020, April. 9). Labour force Survey, MARCH 2020 [Online]. Available: https://www150. statcan.gc.ca/n1/daily-quotidien/200409/dq200409a-eng.htm

[8] Statistics Canada. (2020, June. 30). Gross domestic product by Industry, April 2020 [Online]. Available: https://www150.statcan.gc.ca/n1/daily-quotidien/200630/dq200630a-eng.htm

[9] Statistics Canada. (2021, April. 22). Gross domestic product (GDP) at basic prices, by industry, monthly, growth rates (x 1,000,000) [Online]. Available: https://www150.statcan.gc.ca /t1/tbl1/en/tv.action?pid=3610043402

[10] S. F. Ardabili, A. Mosavi, P. Ghamisi, F. Ferdinand, A. R. Varkonyi-Koczy, U. Reuter, T. Rabczuk, and P. M. Atkinson, ‘‘ COVID-19 Outbreak Prediction with Machine Learning,’’ vol. 13 , no. 10, 2020.

[11] A. Sedaghat, S. Band, A. Mosavi and L. Nadai, "COVID-19 (Coronavirus Disease) Outbreak Prediction Using a Susceptible-Exposed-Symptomatic Infected-Recovered-Super Spreaders-Asymptomatic Infected-Deceased-Critical (SEIR-PADC) Dynamic Model," 2020 IEEE 3rd International Conference and Workshop in Óbuda on Electrical and Power Engineering (CANDO-EPE), 2020, pp. 000275-000282, doi: 10.1109/CANDO-EPE51100.2020.9337775.

[12] C. K. Leung, Y. Chen, S. Shang and D. Deng, "Big Data Science on COVID-19 Data," 2020 IEEE 14th International Conference on Big Data Science and Engineering (BigDataSE), 2020, pp. 14-21, doi: 10.1109/BigDataSE50710.2020.00010.

[13] C. K. Leung, Y. Chen, S. Shang, Y. Wen, C. C. Hryhoruk, D. L. Levesque, N. A. Braun, N. Seth, and P. Jain, ‘‘Data Mining on Open Public Transit Data for Transportation Analytics During Pre-COVID-19 Era and COVID-19 Era,’’ Advances in Intelligent Networking and Collaborative Systems, pp. 133−144, 2020.

[14] S. Tabik et al., "COVIDGR Dataset and COVID-SDNet Methodology for Predicting COVID-19 Based on Chest X-Ray Images," in IEEE Journal of Biomedical and Health Informatics, vol. 24, no. 12, pp. 3595-3605, Dec. 2020, doi: 10.1109/JBHI.2020.3037127.

[15] D. Haritha, N. Swaroop and M. Mounika, "Prediction of COVID-19 Cases Using CNN with X-rays," 2020 5th International Conference on Computing, Communication and Security (ICCCS), 2020, pp. 1-6, doi: 10.1109/ICCCS49678.2020.9276753.

[16] J. Li, D. Zhang, Q. Liu, R. Bu and Q. Wei, "COVID-GATNet: A Deep Learning Framework for Screening of COVID-19 from Chest X-Ray Images," 2020 IEEE 6th International Conference on Computer and Communications (ICCC), 2020, pp. 1897-1902, doi: 10.1109/ICCC51575.2020.9345005.

[17] B. K. Umri, M. Wafa Akhyari and K. Kusrini, "Detection of Covid-19 in Chest X-ray Image using CLAHE and Convolutional Neural Network," 2020 2nd International Conference on Cybernetics and Intelligent System (ICORIS), 2020, pp. 1-5, doi: 10.1109/ICORIS50180.2020.9320806.

[18] A. M. Ibrahim, M. Mohammed Eid, N. N. Mostafa, N. E. -H. Mohamed Bishady and S. H. Elghalban, "Modeling the effect of population density on controlling Covid-19 initial Spread with the use of MATLAB numerical methods and stringency index model," 2020 2nd Novel Intelligent and Leading Emerging Sciences Conference (NILES), 2020, pp. 612-617, doi: 10.1109/NILES50944.2020.9257960.

[19] A. C. Euloge Mouvoh, A. Bouchnita and A. Jebrane, "A contact-structured SEIR model to assess the impact of lockdown measures on the spread of COVID-19 in Morocco’s population," 2020 IEEE 2nd International Conference on Electronics, Control, Optimization and Computer Science (ICECOCS), 2020, pp. 1-4, doi: 10.1109/ICECOCS50124.2020.9314462.

[20] A. Oumina, N. El Makhfi and M. Hamdi, "Control The COVID-19 Pandemic: Face Mask Detection Using Transfer Learning," 2020 IEEE 2nd International Conference on Electronics, Control, Optimization and Computer Science (ICECOCS), 2020, pp. 1-5, doi: 10.1109/ICECOCS50124.2020.9314511.

[21] S. M. Iacus, F. Natale, C. Santamaria, S. Spyratos, and M. Vespe, ‘‘Estimating and projecting air passenger traffic during the COVID-19 coronavirus outbreak and its socio-economic impact,’’ Safety Science, vol. 129, p. 104791, 2020.

[22] P. R. Jena, R. Majhi, R. Kalli, S. Managi, and B. Majhi, ‘‘Impact of COVID-19 on GDP of major economies: Application of the artificial neural network forecaster,’’ Economic Analysis and Policy, vol. 69, pp. 324−339, 2021.

[23] Statistics Canada. Table 36-10-0434-01 Gross domestic product (GDP) at basic prices, by industry, monthly (x 1,000,000) doi: https://doi.org/10.25318/3610043401-eng

[24] D. D.Béland and A. Waddan, ‘‘Social policy and the recent economic crisis in Canada and the United States,’’ Social policy in challenging times, pp. 231−250.

[25] Our World in Data. Table Explore the global data on confirmed COVID-19 cases Available: https://ourworldindata.org/covid-cases

[26] E. Goddard, ‘‘The impact of COVID‐19 on food retail and food service in Canada: Preliminary assessment,’’ Canadian Journal of Agricultural Economics/Revue canadienne d'agroeconomie, vol. 68, no. 2, pp. 157−161, April 2020.

[27] V. Singh, Economic contribution of culture in Canada. Ottawa: Statistics Canada, Culture, Tourism & the Centre for Education Statistics, 2004.

[28] M. Watts, L. E. Ritchie, J. R. Brown, M. Fekete, J. Ranger, S. Newell, M. Putyra, M. Longo, P. Sullivan, P. Olexiuk, C. Wetter. (2020, Mar. 24). Provincial governments restrict business operations in the fight against COVID-19 [Online]. Available: https://www.osler.com/PDFs/Resource/en-ca/COVID-19-Ontario-and-Quebec-Governments-Limit-Non.pdf

[29] D. J. Andrea and B. C. Smith. The Canada-US Border: An Automotive Case Study. Center for Automotive Research, 2002.

[30] E. Mazareanu, ‘‘Flights cancellation rate of global airlines due to COVID-19 as of March 23, 2020’’ Statista [Online], Jun 18 2020. Available: https://www.statista.com/statistics/1111989/ flights-cancelled-airlines-worldwide-covid-19/

[31] R. Barichello, ‘‘The COVID‐19 pandemic: Anticipating its effects on Canada's agricultural trade,’’ Canadian Journal of Agricultural Economics/Revue canadienne d'agroeconomie, vol. 68, no. 2, pp. 219−224, 2020.

[32] Statistics Canada. (2020, April. 30). Gross domestic product by industry, February 2020 [Online]. Available: https://www150.statcan.gc.ca/n1/daily-quotidien/200430/dq200430a-eng.htm

[33] R. Madurai Elavarasan, G. M. Shafiullah, K. Raju, V. Mudgal, M. T. Arif, T. Jamal, S. Subramanian, V. S. Sriraja Balaguru, K. S. Reddy, and U. Subramaniam, ‘‘ COVID-19: Impact analysis and recommendations for power sector operation,’’ Applied Energy, vol. 279, p. 115739, 2020.