Measurable Impacts of the Covid-19 Pandemic on Different Sectors of the Global Economy

Project Report

Aladin Djuhera 03671760
Michael Brandner 03681315
Alexander Griessel 03687564
Maximilian Putz 03682862
Niklas Landerer 03671815
Felix Montnacher 03689999
Martin Schuck 03683571
Henrique Soares Frutuoso 0368213
Aron Endres 03689152

September 4, 2020

Motivation

Recently, the Covid-19 pandemic and its consequences have dominated public discourse and scientific research. It is yet unclear to what extent the sheer scale of the Covid-19 crisis will impact our society, our economy, and the environment. Restrictions imposed by governments as well as divergent personal behaviours of citizens have changed the way our society operates. Due to social distancing policies inducing an increased demand for virtual replacements of social and workplace interactions, internet traffic has surged. Meanwhile, travel and manufacturing industries have taken a substantial blow from lock-down measures.

The quantitative extent the impact the pandemic has had or still has on different industries is a major subject of public debate. The common narrative sees tech companies as the determinate winners of the pandemic months, and conventional industries as the losers. We deem the verification and quantification of these statements to be an important contribution to public debate. Furthermore, the quantitative analysis of the impact should be of forefront interest for the public and for policy-makers.

As a society, we must decide which branches of our economy we deem necessary to prioritise for the allocation of limited subsidisation funds during the impending threat of a global recession. The extent of damages suffered due to the Covid-19 outbreak must be recognised if the public debate is to remain impartial and fair.

---

1 Project Description

This project aims to quantify the impact that the Covid-19 pandemic will have on the global economy, utilising machine learning. One of the main limitations in machine learning is the modelling of unsampled regions in the feature space. Furthermore, models learned during the training process are only valid in the context of unchanged underlying data generation processes. Since the Covid-19 pandemic fundamentally changed the way our society operates, and we don't have any samples of a global pandemic of smaller and/or greater size, it is impossible to directly model the impact of the virus on different economical sectors. Instead, we aim to model the economy in absence of the Covid-19 pandemic, under the assumption that the pandemic began shortly after early January. The impact of the pandemic can then be measured as the integral of the difference between the predicted courses without Covid-19, and the actual observed outcomes. The resultant models and their predictions are made publicly accessible with an interactive web interface.

Research Question

To concisely formulate the scope of our project, we pose the following research question: can we measure a significant difference in key indicators of internet activity and different economic branches between the ground truth progression and the predicted pandemic-absent progression? We lay a special focus on the technology sector, as it is commonly assumed that this sector is one of the few winners of the pandemic- due to an alleged surge in internet activity.

Goals

In this project, we aim to gather relevant data about both the economy and online activity during and before the Corona pandemic. The data is to be analysed and processed in a data pipeline. Furthermore, a model selection process shall determine suitable machine learning estimators and their respective hyper-parameters for predicting the progression of the various indicator time series without the influence of the pandemic. We then want to analyse and quantify if a surge in internet activity has indeed occurred, and if so, determine whether it had a dominant influence on the economy. We especially want to answer if the common assumption that the tech sector was least affected by the pandemic due to this increased use of technology holds up against our predictions.

2 Data basis

The data we use in this project can be separated in two distinct groups.
The first group consists of key indicators related to internet usage and consumption. These indicators include statistics on view counts and streamed hours of social media platforms, but also metrics like network traffic through the largest internet exchange points. To our astonishment, these statistics are not as accessible as we initially thought. Although, or probably because, tracking data of users has become one of the most valuable resources of internet services, companies are reluctant to disclose any of their collected data.
The second data group focuses on finance, which is publicly available and can easily be obtained from the Alphavantage Finance API.

Sources

All finance data is obtained by calls to the Alphavantage Finance API. The data we pulled covers a large number of the world's biggest publicly traded companies (for their respective fields). Data comprises of these companies' stock performances in the past few years.

Data for internet behaviour stems from numerous sources. For gaming, we use the unofficial Steam and PlayStation statistics databases steamdb.info and gamstats.com. Unfortunately, Microsoft does not publish any data on the Xbox Live Network. Streaming data originates from twitchtracker.com. Again, big streaming providers such as Netflix or Amazon Prime do not provide detailed statistics on their user habits. The data on YouTube views comes from the socialblade.com website, a third party YouTube statistics website. Since API access to Socialblade is paywalled, we employ a custom scraper further described in the next section to deal. Finally, for a rough estimation of worldwide internet traffic we use data from several Internet Exchange Points (IX). Since each Internet Exchange Point offers access to its usage statistics in a custom way, a mixture of API calls and graph readers is used.

Data collection

Accessing Steam and PlayStation data is a matter of visiting their respective unofficial web-pages, and downloading a .csv file containing the relevant statistics. Data is then manually stored within our repository.

The twitchtracker website does not offer a file download, or an API. However, since all of the data is contained within the HTML file of the landing page's URL, we can simply extract Twitch data by processing the content of a single request to the landing page.

YouTube is the most dominant video platform on the Internet. Nevertheless, accurate information on its user statistics is not publicly available. We use the third party tracking website Socialblade instead. Socialblade tracks weekly activity statistics for the most influential channels on YouTube. Since Socialblade only offers the statistics for the top 250 subscribed channels per country, we use those channels as an approximation for the overall traffic on YouTube. Combined with more than 240 available countries, we can track data for more than 50.000 of the most influential channels on Youtube. It should be noted that this indicator relies on some strong assumptions. First, top channels do not currently have to include sub-channels that were successful in 2018. Also, traffic increases in minor channels, even when significant, are not tracked by this method. As such, data is gathered under the assumption that changes in the top 250 channels by country are representative of changes in minor channels.

Socialblade scraping is accomplished in three steps:

- Fetch relevant countries and their channels from Socialblade.
- For each country, get the list of the top 250 subscribed channels.
- For each of the resulting channels, scrape all available data from the stats web-page.

This process poses several challenges. Since Socialblade is protected against DDoS attacks and bots by Cloudflare, we needed to develop a custom scraper. Scraping the complete list of channels requires approximately 55.000 requests to the website, which results in a Cloudflare ban. We employ VPN connections that change the exit server once Cloudflare blocks access. To handle the workload, the scraping process divides the work into smaller packages. These packages can then be worked on by multiple scrapers running as docker containers on the same machine. Figure 1 shows the scraping in process.

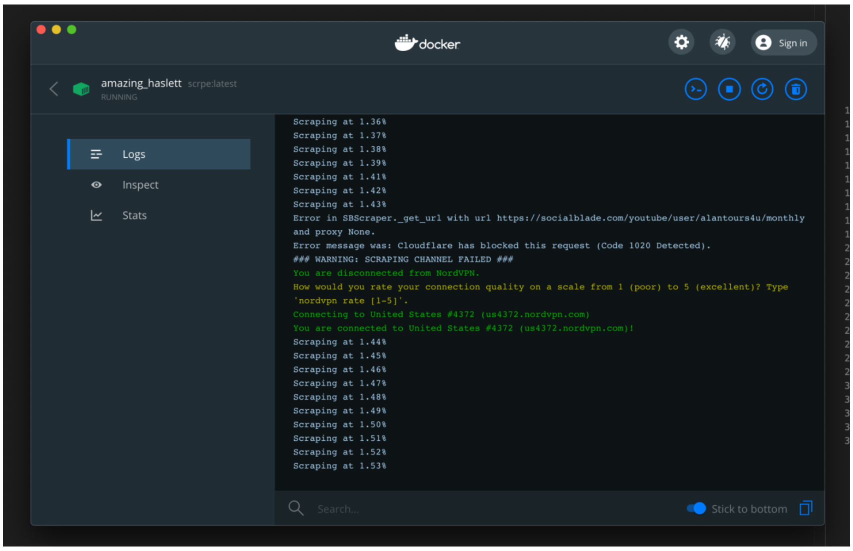

Figure 1: Scraping process in action.

To obtain the IX data set we employ a mix of conventional API calls and automatic graph readers. Although the DE-CIX does provide an API to a vast collection of the worlds largest Internet Exchange Points' statistics, access is unfortunately blocked by a paywall. An inquiry to gain unpaid access for academic purposes did not yield any response. Instead, we opted for a manual collection of the data using a list of the most commonly used IX points and their respective websites. [Link]. The non-homogeneous data representation from each IX necessitates a mix of strategies. For IX points that do not offer a direct .csv download, we use a scraper that has to be adapted for each website. In some cases, statistics are only provided as pictures. In these cases, we use a graph reader [Link] with some manual labour to create a usable .csv file.

The financial data is made up of the stock market values of the most important companies in each of their respective sectors. The sectors covered are: medical, financial, energy, oil, steel, automotive, telecommunications, tech and stock market indices. A detailed list of the sectors as well as the companies belonging to them can be found in the appendix. We normalise the stock data to zero mean and unit variance before averaging over them in order to give the companies equal influence in the branch performance time series. As a result, we obtain a more abstract, unit-less performance indicator for each sector, which is better suited to reflect the relative tendency of the whole sector and not only the trend of the highest value shares.

Preprocessing

As our data stems from many different locations, each source has its own challenges. We use a single pipeline which requires processing functions for each data set to follow a unified interface. Not all preprocessing steps should be performed on all data sets. For example, Twitch statistics are given on a monthly basis and are interpolated to match the weekly sampling frequency of other data sets. Other data sets such as IX traffic are prone to excessive noise. For these series we employ rolling averages over the data to deal with jitter. The soundness of our preprocessing steps are verified by visual inspection of the data plots. For some series such as the IX and Socialblade data, summation and re-binning of many samples is necessary. This is done during the raw data processing *prior* to entering the pipeline. The only exception is the finance data, where the pipeline averages over all normalised stock performances of a domain to form a single performance indicator for each sector.

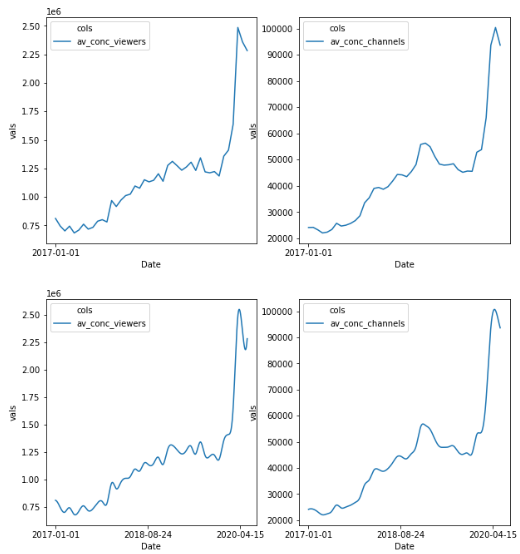

Figure 2: Example of the pipeline processing. Time series with insufficient sampling rates are interpolated to a fitting density. Visualisation ensures sound data interpolation.

After each dataset is processed individually, all numerical categories are scaled with a standard scaler to avoid any numerical problems in respect to our models. We do not make use of dimensional reductions such as PCA/PLS. Our project relies on the meaningful prediction of single time series, which renders these approaches completely unfitting for our data preprocessing.

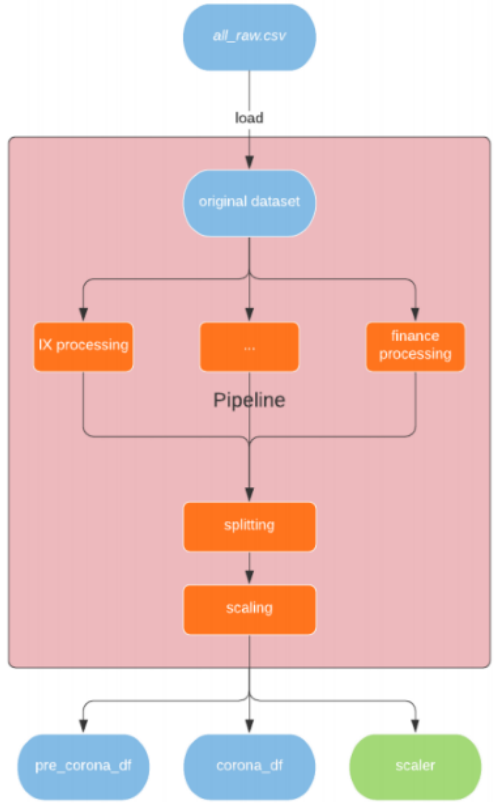

Figure 3: Float chart of the pipeline.

3 Data Model

When dealing with the Corona pandemic in the context of machine learning, we face two main challenges. Firstly, the pandemic fundamentally changed some of the underlying data generating mechanisms of our society. Machine learning models heavily rely on distributions not shifting too heavily, as they are designed to capture the particular distribution during training. A change in those mechanisms renders the model's predictions inaccurate. Secondly, most of the interesting questions regarding Covid-19 are concerned with the course of future events, or the hypothetical courses of events under different conditions. These questions require the modelling of time series, a very delicate topic in machine learning. Although not impossible, one has to be especially careful with these kinds of predictions. There are several approaches to time series predictions such as Gaussian Processes or Recurrent Neural Networks. In our project, we employ an individual additive decomposition model with trend and seasonality estimation for each time series in our data set. We make use of a Python implementation of this approach from Facebook, namely Facebook Prophet [Link], which uses partial Fourier sums for seasonality estimation [Link].

Approach

We approach the problem by splitting each time series into pre-Corona and Corona time-frames. We chose January 1st as the starting date of the pandemic, based on the number of global cases and spread of the virus. Although one could argue February to be a more suitable starting date, the effect on our models should be negligible as they are designed to accurately predict the development of series yet to be influenced by the pandemic. Using the pre-Corona data, a time series predictor is trained to predict the hypothetical course of the time series without the knowledge of a global pandemic. This approach yields forecasts for each series, which are then compared to the data that was observed during the pandemic. The deviations of the courses are then interpreted as the impact of the pandemic on economic and online activities. The obvious assumptions and limitations of this approach will be discussed later. As mentioned before, the main motivation with predicting a hypothetical world without a pandemic is that models based on pre-Corona data can't account for mechanisms encountered during the pandemic. The only way of making a sound statement about Corona induced changes is to model a world without the Covid-19 crisis, and compare the results with current observations.

Training

For the model training we split each time series into a training and test set. We perform grid-search for the optimal hyper-parameters by using 5-fold cross-validation on our candidate estimators. Grid-search was performed on three estimators: an Extreme Learning Machine, a Gaussian Process, and an Additive Decomposition estimator. To rank our estimators, we use the R-2 score as the error metric for our test sets. In order to unify the cross-validation process across different candidate estimators, we employ a custom interface to Facebook Prophet that exposes the same functionality as the sklearn Basis Estimator. This enables cross-validation of all models under a common framework. Due to the low computational complexity of our candidate models, grid search is an option, unlike for more elaborate and complex models, such as Recurrent Neural Networks.

Evaluation

We evaluate the performance of each model on a separate hold-out test set for each time series. The test set, just as the cross-validation sets, consists of a timeframe taken from the time series. Sampling the test set at random from the whole data set would yield misleading results because of the strong correlations between successive data points. Out of the three models, the Additive Decomposition estimator establishes itself as the model with the highest performance. It should again be emphasised that cross-validation for hyper-parameter search and model testing is performed on every time series individually, and that the decomposition approach outperforms other estimators in every case. In fact, Gaussian Processes and ELMs prove to be barely able to produce reasonable forecasts, while the additive decomposition not only outperforms in terms of error metrics, but also provides predictions that prove to be visually convincing.

4 Results

Following our data set nature, our observations can be divided into two parts. Firstly, we try to verify that the Corona pandemic has indeed lead to a significant increase in internet traffic. Secondly, we show that at the start of the pandemic, economic sectors across the board take a massive hit in their average economic performance and observe their individual recovery time spans.

Observations

Our models show a significant rise in online activity across the board. All predicted courses of key indicators deviate significantly from the measured data once the Corona pandemic sets in. Raw data traffic, video views on YouTube, gaming activity on some of the most popular networks (steam and PlayStation Network) and streaming on Twitch increased well beyond the predicted levels without the corona pandemic. Figure 4, 5 and 6 show representative plots of our predictions. A complete list of all predictions can be found in appendix 8. Internet traffic in bit/s across our scraped IX points at peak for example rose by 12.4% compared to the base line our model predicted. Twitch average views were up by a maximum of 101%, and the PlayStation Network had an increase of 236% at the largest discrepancy.

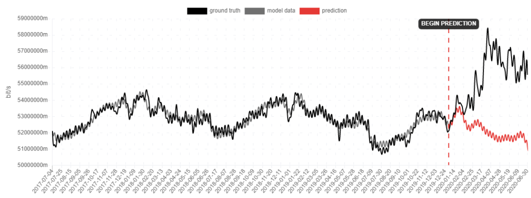

Figure 4: Predicted vs actual progression of the summed bit rate of a collection of the worlds largest IX points.

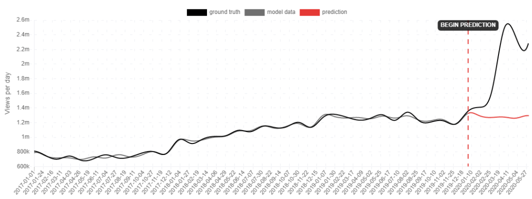

Figure 5: Predicted vs actual progression of the average views per week on one of the world's largest streaming platforms, Twitch.

We furthermore observe a massive decline in the averaged economic performance across all branches as soon as the pandemic takes effect across the globe. Some indices recover astonishingly fast, up until the point of reaching levels that our model predicted for a progression without any pandemic at all. Other sectors on the other hand continue to struggle with the consequences of the pandemic. Figure 4, 5 and 6 show representative plots of our economic predictions. As mentioned before, the complete list of all predictions can be found in appendix 8. Counter to our initial expectations, some of the traditional industry sectors such as steel seem to be less affected

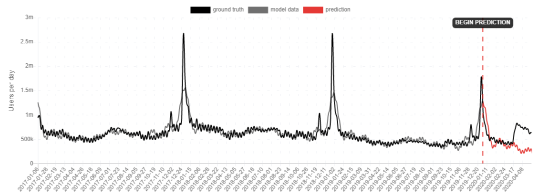

Figure 6: Predicted vs actual progression of active users on the PlayStation Network.

than technology. Although the technology sector recovers remarkably well, so do several traditional branches- such as the automotive sector. We created an abstract performance index for each sector as described in the preprocessing section. This index is of 0 mean and unit variance. Instead of dollars, we therefore report discrepancies in terms of standard deviations to get a meaningful comparison between the different indices. Prediction and ground truth data for the oil sector for example differ by 2.44 standard deviations at its largest point. Furthermore, the performance does not recover significantly until the end of our prediction window (30.6.2020). Tech had a discrepancy of 1.54 standard deviations, but closes at 0.13 deviations. This marks a near complete recovery of the tech branch. The finance sector had its peak difference at 2.42 standard deviations and does not recover, just like the oil sector.

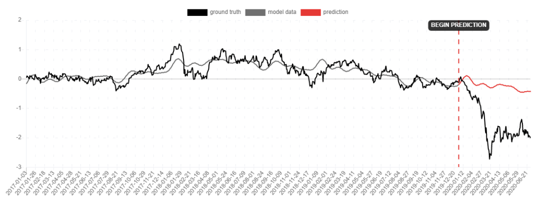

Figure 7: Predicted vs actual economic performance of a collection of the world's largest oil sector companies.

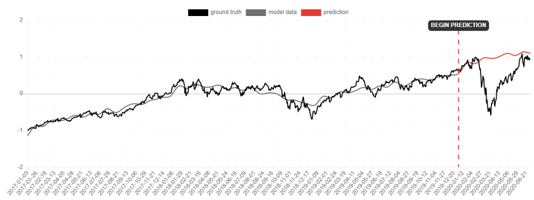

Figure 8: Predicted vs actual economic performance of a collection of the world's largest tech sector companies.

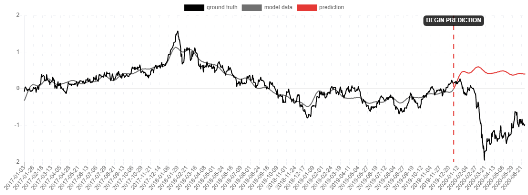

Figure 9: Predicted vs actual economic performance of a collection of the world's largest finance sector companies.

To our big surprise, the steel sector seems to be mostly unaffected by the pandemic. At its worst point, it slumped by 0.64 standard deviations and closed at 0.21 deviations.

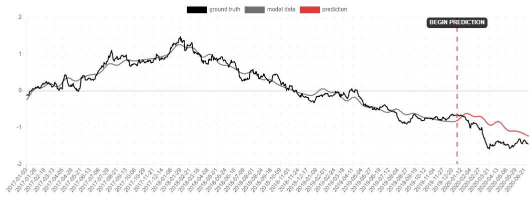

Figure 10: Predicted vs actual economic performance of a collection of the world's largest finance sector companies.

We compose a ranking of all branches, sorted from the least affected branch at the end of our prediction window to the worst. The impact, as before, is measured in standard deviations of the economic index.

1. Technology: 0.13
2. Steel: 0.22
3. Medical 0.73
4. Automotive 0.74
5. Energy 0.97
6. Telecom 1.16
7. Finance 1.37
8. Oil 1.59

Trends

Initially, our models match the progression of most time series pretty well. However, it is clear that all key indicators of internet activity see a sharp incline at the time Covid-19 starts to spread around the globe compared to what our model predicts as a "normal" progression under non-Corona conditions. Similarly, the average economic performance across all sectors drastically declines at the same point in time. Apart from the initial slump, another interesting trend observation is that a number of indicators for both internet activity and economic performance seem to approach the levels our model predicts for a course without any Corona pandemic.

5 Discussion

The following section presents a possible interpretation of the observed results. It should be noted that our observations show a strong correlation between the Corona pandemic, the rise of internet activity and the sharp decline across nearly all economic sectors. This correlation must not be mistaken for causality! We do try to give a causal interpretation of the results based on our reasoning. This does not mean however that a direct or indirect causal connection exists between our results.

Interpretation of the results

The discrepancies between our modelled courses and the actual observations strongly suggest the pandemic had a major impact on internet consumption by increasing traffic across a wide variety of platforms, as well as absolute data traffic. We interpret this as a consequence of government enforced lockdowns around the world. The lockdowns forced people to increasingly rely on home solutions for work and entertainment, which directly correlates with the overall traffic and platform specific loads. As a further consequence of these lockdowns, all economic sectors we look at took a huge blow. We attribute this to the fear of the economic impact of the pandemic spreading across the markets. Surprisingly, some traditional industries like the steel sector were less affected than we initially thought. We are unsure as to why that is, but one reason could be that branches not directly related to the consumer market are less bound to immediate consumer decisions, and are more likely to be bound to delivery contracts in the industrial sector. Also, counter to the common narrative, the tech sector wasn't the only sector to make a speedy recovery once the initial shock was over. We take this as an indication that, although certainly a contributing factor, the shift towards online communication and entertainment is not the main driver constituting a rapid rehabilitation.

Critical assessment of the results and assumptions

Our approach relies on several strong assumptions. For example, we postulate that the stock market is a depiction of current economic development. We further assume that both economic and internet activity are suitable for a trend/seasonality/noise decomposition. Some series, such as traffic on the PlayStation Network, fit this framework seamlessly. Others, like the medical sector's performance, seem to be strongly governed by external and non-periodic events, which leads to a significant performance decrease of our decomposition model. Another big assumption is that the difference between the predicted and observed courses of our time series is caused by the Covid-19 pandemic. Although intuitive, this does not necessarily have to be the case. One could imagine a scenario where one branch is heavily affected by a technological breakthrough, a political decision or some other event not related to the pandemic at all. Such events couldn't possibly be predicted by an estimator that only has access to the time series. However, the extent to which the Corona crisis has influenced all parts of our society is so thorough that during the last 6 months, nearly no event of significance that is completely unrelated to Covid-19 has taken place- at least not event as noteworthy. We therefore deem this assumption to be largely satisfied.

As with all time series predictions, the confidence in the ability of our models to accurately predict a hypothetical course without a global pandemic declines with the size of the time interval we try to predict. Since we train on 2-6 years of data for each time series, predicting a few months seems to lie within a reasonable time frame. However, we do note that this approach is not suited for a long-time analysis of the impact Corona has had on both the internet and the economy. Predictions on the future, especially when dealing with series as highly unstable as the stock performance of different branches, should be handled with healthy scepticism.

Nevertheless, the results do provide a quantitative estimation on the impact of Corona on the change.

Proposed Answer to the Research Question

We observe the suspected surge in internet activity when comparing our model projections to the ground truth data. Likewise, the ground truth data shows a sharp decline in the economy across nearly all sectors compared to our model predictions. The answer to our original questions is therefore a yes: we can indeed measure a significant difference in both categories. Whether the rise in internet activity has had a major influence on the technology sector's speedy recovery could not be fully verified. There is undoubtedly a correlation between both progressions, but it need not be causal.

6 Conclusion

During our project, we were able to answer the research question posed at the very beginning. We started by assembling a data basis for the project from various sources. To this end, we used readily available data sets, API call scripts as well as more complicated custom web scrapers. We then developed a data processing pipeline and implemented a method to guide our model choice by cross-validation and testing on a hold out test set. The models were trained to predict the time series of our data sets during the times prior to Corona. Once we obtained models we were confident in, we predicted the course of our time series for the last six months to get a projection on what the data would have looked like without a global pandemic. The comparison of the ground truth data during the pandemic and the modelled data was then measured and interpreted as the impact of the Covid-19 crisis.

Summary of the Results

We showed that a significant rise in internet activity ensued during the Covid-19 pandemic by comparing the observed ground truth data of several key indicators to data projected by models trained on the past time series before the Corona crisis. Furthermore, we analysed the global economy separated into sectors with the same procedure. We used an average of the normalised stock performance of the branches biggest companies to get a meaningful indicator of the branches economic well-being. Comparisons to our projected courses show a sharp decline around the time the lockdowns ensued around the world. We delivered an approximate quantification for the impact on both topics. We weren't able to confirm the common assumption that the technology sector is the least affected sector due to the previously mentioned surge in internet usage.

Future Work

One of the biggest challenges of this project was the collection of suitable data. In the future, a more thorough collection of data sets could shed more light on the impact in different parts of everyday life, or corroborate our findings. However, time series predictions become more and more unreliable the more time passes. Therefore, there is a limited time frame to use our method. Otherwise, in the unfortunate event of similar ensuing pandemics in the future, one could try to model economic impact and overall human behaviour based on the severeness of the pandemic. Using the data from just one global pandemic however is not enough to train sound models for this task.

7 Comments to the Group Work Experience

Working in a group of nine people was a challenge in terms of logistics, organisation, and work load. First, a fitting organisational structure had to be agreed on. Oftentimes, we found a rigid role assignment to be impractical. Team members working on a particular sub task are much more suitable to oversee this part of the project than an external committee. Nevertheless, one of the most important things we learned was that clear responsibilities are essential for a large group to be functional. Once we divided into a web development and machine learning group, things became much more streamlined. It also took some time to accurately access the strengths and skills of all team members and where to apply them best. Although we faced some high pressure situations and major setbacks, overall the group remained in good spirit and harmonised well. We feel like there is a large potential to further streamline the communication.

In terms of communication, we felt that the team was most productive when the majority of meetings occurred between team members entrusted with the same or similar tasks. However, regular and well-scheduled meetings with complete attendance were an important factor in retaining a level of oversight for each team member. The less organised and granular the meeting structure was, the longer meetings became in general, as overarching meetings were used to discuss minor or specific problems largely irrelevant to the entire team.

Finally, we found that structuring a team of students actively participating in university classes could not be accomplished by running the team like a project team at a regular company. Differing schedules, available time, and differing cumulative experience meant that team structure was largely fluid and had to remain adaptable.

In the end, our team worked well together, with an end result the team is happy with, and relatively even average workloads.

8 Appendix

Complete collection of predicted time series plots

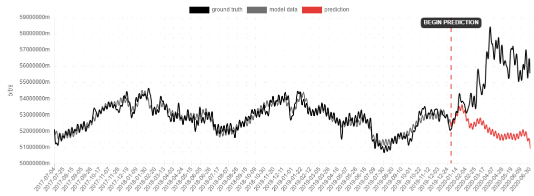

Figure 11: Predicted vs actual progression of the summed bit rate of a collection of the worlds largest IX points.

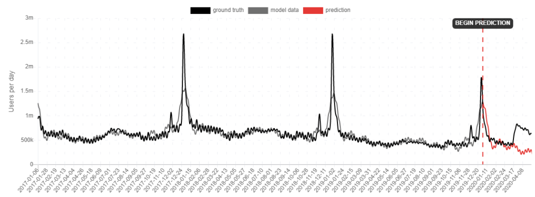

Figure 12: Predicted vs actual number of active accounts on the PS Network.

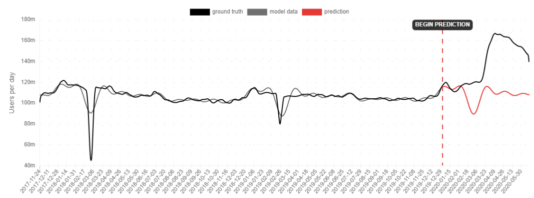

Figure 13: Predicted vs actual number of active accounts on steam.

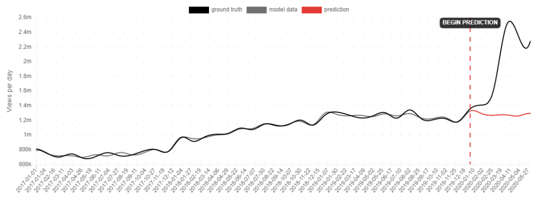

Figure 14: Predicted vs actual progression of the average views per week on one of the world's largest streaming platforms, Twitch.

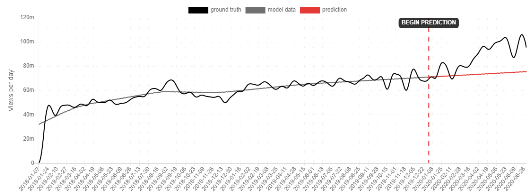

Figure 15: Predicted vs actual number of viewers over a selection of the worlds most subscribed channels.

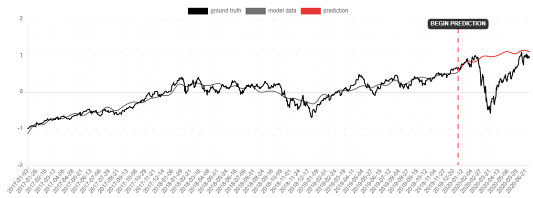

Figure 16: Predicted vs actual economic performance of a collection of the world's largest tech sector companies.

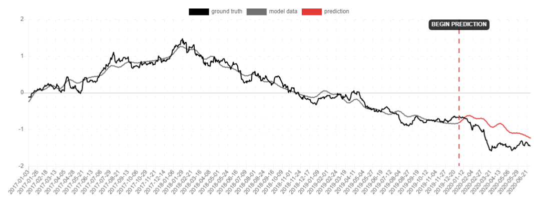

Figure 17: Predicted vs actual economic performance of a collection of the world's largest steel sector companies.

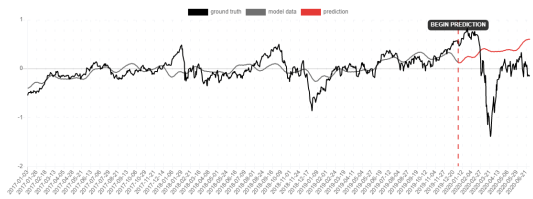

Figure 18: Predicted vs actual economic performance of a collection of the world's largest medical sector companies.

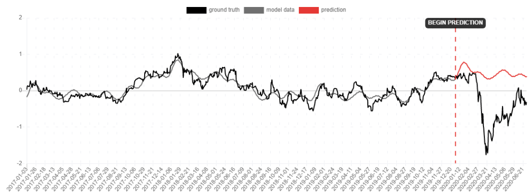

Figure 19: Predicted vs actual economic performance of a collection of the world's largest automotive sector companies.

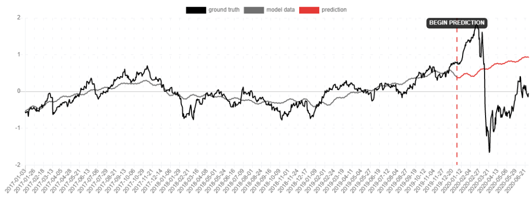

Figure 20: Predicted vs actual economic performance of a collection of the world's largest energy sector companies.

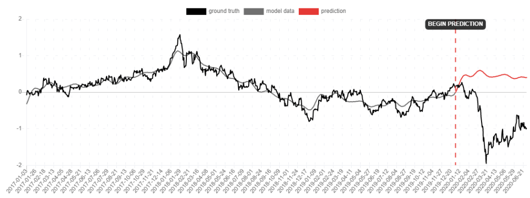

Figure 21: Predicted vs actual economic performance of a collection of the world's largest finance sector companies.

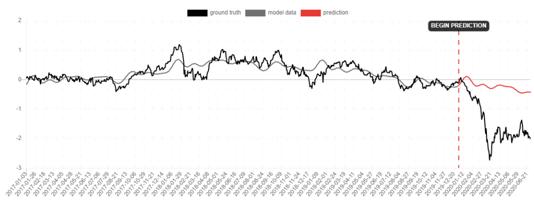

Figure 22: Predicted vs actual economic performance of a collection of the world's largest oil sector companies.

Finance data

Medical companies

Symbol	Company
EVT	Evotec AG
SHL	Siemens Healthineers AG
BAS	BASF
BAYN	Bayer AG
FMS	Fresenius Medical
FRE.DE	Fresenius SE
JNJ	Johnson + Johnson
PFE	Pfizer
ABT	Abbott Laboratories
KMB	Kimberly Clark Corporation
MDT	Medtronic PLC
PHG	Philips Healthcare
GE	General Electric (Healthcare)
BDX	Becton Dickinson
CAH	Cardinal Health
SYK	Stryker Corporation

Table 1: Medical companies with associated stock exchange abbreviation

Financial companies

Symbol	Company
DB	Deutsche Bank
ICK	Industrial and Commercial Bank of China
GS	Goldman Sachs
CMC	JPMorgan Chase
BRYN	Berkshire Hathaway Inc.
NCB	Bank of America
NWT	Wells Fargo and Company
JPHLF	Japan Post Holdings Co. Ltd.
CICHY	China Construction Bank Corporation
ACGBY	Agricultural Bank of China
CRARY	Credit Agricole SA
BACHF	Bank of China Ltd.
C	Citigroup Inc.

Table 2: Financial companies with associated stock exchange abbreviation

Stock indices

Symbol	Company
DAX	DAX
TDXP	Tech Dax
INDU	Dow Jones Industrial
SNP	SandP 500
NDAQ	NASDAQ
NKY	Nikkei 225

Table 3: Stock indices with associated stock exchange abbreviation

Energy companies

Symbol	Company
SIEGY	Siemens AG
EOAN.DE	E.ON
RWE.DE	RWE AG
DUK	Duke Energy Corporation
ENGI.PA	Engie SA
NGG	National Grid PLC
NEE	NextEra Energy Inc.
EDF	Electricite de France SA

Table 4: Energy companies with associated stock exchange abbreviation

Oil companies

Symbol	Company
CVX	Chevron Corporation
XOM	ExxonMobil
PTR	PetroChina
RDS-A	Royal Dutch Shell
LUK	Lukoil
ROSN	ROSNEFT
TOT	Total SE
BP	BP Plc
SNP	China Petroleum and Chemical Corporation

Table 5: Oil companies with associated stock exchange abbreviation

Steel companies

Symbol	Company
TKA.DE	ThyssenKrupp AG
MT	ArcelorMittal S.A.
NISTF	Nippon Steel and Sumitomo Metal Corporation
SHE:000709	Hebei Iron and Steel Group
SHA:600019	Baoshan Iron & Steel (Baosteel) Co. Ltd.
PKX	Pohang Iron and Steel Company (Posco)
SHE:002075	Jiangsu Shagang

Table 6: Steel companies with associated stock exchange abbreviation

Automotive companies

Symbol	Company
TM	Toyota Motor Corporation
GM	General Motors Company
HYMTF	Hyundai Motor Company
BMW.DE	BMW AG
NSU	AUDI AG
VOW.DE	Volkswagen
DAI.DE	Daimler AG
CON	Continental AG
TSLA	Tesla Inc.

Table 7: Automotive companies with associated stock exchange abbreviation

Telecommunication companies

Symbol	Company
DTE	Deutsche Telekom AG
DRI	1&1 Drillisch AG
TEF	Telefonica SA
O2D.DE	O2 Deutschland AG
T	AT&T Inc.
TMUS	T-Mobile US Inc.
VOD	Vodafone Group
CTM	China Mobile Ltd.
VZ	Verizon Communications
NTT.F	Nippon Telegraph & Tel
SFTBY	Softbank Group Corporation
AMOV	America Movil
CHA	China Telecom

Table 8: Telecommunication companies with associated stock exchange abbreviation

Tech companies

Symbol	Company
AAPL	Apple Inc.
AMZN	Amazon.com Inc.
GOOGL	Aplhabet Inc.
CCCMF	Cancom SE
IFX.DE	Infineon AG
SAP	SAP SE
CSCO	Cisco Systems Inc.
IBM	IBM
INL	Intel Corporation (DE)
INTC	Intel Corporation
MSF	Microsoft Corporation
EBAY	eBay Inc.
EA	Electronic Arts Inc.
TWTR	Twitter Inc.
QCOM	Qualcomm Inc.
SNE	Sony Corporation
TXN	Texas Instruments Inc.
QCOM	Qualcomm Inc.
NFC	Netflix Inc.
ZM	Zoom Video Communications Inc.

Table 9: Tech companies with associated stock exchange abbreviation