2022 44th Annual International Conference of
the IEEE Engineering in Medicine & Biology Society (EMBC)
Scottish Event Campus, Glasgow, UK, July 11-15, 2022

2022 44th Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC) | 978-1-7281-2782-8/22/$31.00 ©2022 IEEE | DOI: 10.1109/EMBC48229.2022.9871146

An approach to a wrist wearable based Covid-19 prediction system to protect Health Care Professionals

Dibyanshu Jaiswal[1], Dr. Kayapanda Mandana[2], Ramesh Kumar Ramakrishnan[1], Kartik Murlidharan[1], Mithun Basaralu Sheshachala[1], Shakil Ahmad[1], Tanmay Acharia[2], Loknath Tiwari[2], Arpan Pal[1] and Balakumar Kanagasabapathy[1]

***Abstract*— With healthcare professionals being the frontline warriors in battling the Covid pandemic, their risk of exposure to the virus is extremely high. We present our experience in building a system, aimed at monitoring the physiology of these professionals 24/7, with the hope of providing timely detection of infection and thereby better care. We discuss a machine learning approach and model using signals from a wrist wearable device to predict infection at a very early stage. In a double-blind test on a small group of patients, our model could successfully identify the infected versus non-infected cases with near 100% accuracy. We also discuss some of the challenges we faced, both technical and non-technical, to get this system off the ground as well as offer some suggestions that could help tackle these hurdles.**

I. Introduction

Currently there is one doctor for every 1,445 Indians as per the country's current population estimate of 1.35 billion, which is lower than the WHO's prescribed norm of one doctor for 1,000 people [1]. This shortage of doctors (and nurses) coupled with the fact that frontline health care workers are most likely to be exposed to the virus than anybody else, makes protecting them critical.

Most of the prior art revolve around detection fo Covid-19 from radiology images such as X-rays of lungs via image processing techniques and deep learning models [2]–[4]. But such methods can prove successfull, only when subjects opt for such test post a doctor consultation. Recently a study using two years of Fitbit data from nearly 50,000 users found that wearable data could more accurately predict local flu outbreaks than the standard system used by the Centers for Disease Control and Prevention [5]. Other works such as [6], [7], show how self reporting and wearable data can act complementary to virus testing. Since coronavirus and the seasonal flu share some symptoms, it could be possible to develop an algorithm that utilizes physiological data in building an artifical intelligence (AI) based early warning system. Such a system could help predict the onset of symptoms and identify whether a health care worker might have been infected with the virus (and thus assist in timely action). Also it would give confidence to frontline health care professionals (HCP) that they are healthy and continuously being monitored for risk.

[1]Tata Consultancy Services Ltd. India, Email: dibyanshu.jaiswal@tcs.com
[2]Fortis Healthcare, Kolkata, India

Our approach was to develop an AI based system that will provide an early warning to the health care professionals who are actively involved in the Covid-19 wards and ICUs. The system will monitor the vital parameters such as Heart Rate, Respiratory Rate, Activity, Skin Temperature on a continuous basis with the facility to visualize the trends at individual level. We also propose to collect continuous photo plethysmography (PPG) signals and develop an AI model which can be use to predict early signs of infection.

We teamed up with Fortis Healthcare Limited, a leading integrated healthcare services provider in India that had an active Covid-19 ward to conduct a study involving HCP involved in management of Covid-19. The study was registered as a cross sectional type observational trial at the Clinical Trials Registry-India (CTRI)[1], approved by the ethical committee at Fortis and conducted on volunteers in the age group of 19 to 60. All the volunteers signed an informed consent before participation.

However, given the real world nature of the deployment and the criticality of the study we faced certain challenges both in the technical and non-technical realm. Since HCP use PPE kits, they cannot carry any personal devices like mobile phones or subject themselves to frequent data capture using standard clinical measurement devices like thermometer, pulse oximeter. Hence we used a wrist wearable which could locally store the data and at the end of the day download the data from the device into computer for analysis. Also, given that any Wi-Fi or Bluetooth based data transmissions could interfere with ICU Electronics, any real time data transmission had to be avoided. Our key contribution is an AI based system that can predict the onset of an infection a few days in advance of the time when actual clinical symptoms/complications begin.

II. Deployment Architecture

With the onset of a viral infection like the seasonal flu, measures such as temperature and resting heart rate often rise slightly. Further, as fatigue settles in, activity levels drop and time spent sleeping rises. It therefore makes sense to track these measures continuously.

Most modern fitness trackers and smart watches such as the Fitbit or the Samsung Galaxy Gear include sensors such

[1]Clinical Trials Registry- India (www.ctri.nic.in), Identifier: CTRI/2020/05/025495

TABLE I: Sensor-Measure mapping

Sensor	Derived Measure
PPG	Hourly heart rate (HR) - Estimated instantaneous HR averaged over the hour [9].
PPG	Hourly resting heart rate - Average heart rate of 5 minutes post a 10 minute rest detection.
PPG	Hourly respiratory rate (RR) - Estimated instantaneous respiratory rate averaged over the hour [9].
Infrared Thermopile	Average skin temperature for the hour.
IMU	Cumulative step-count of the hour [10].
EDA Sensor	To detect watch removal events.

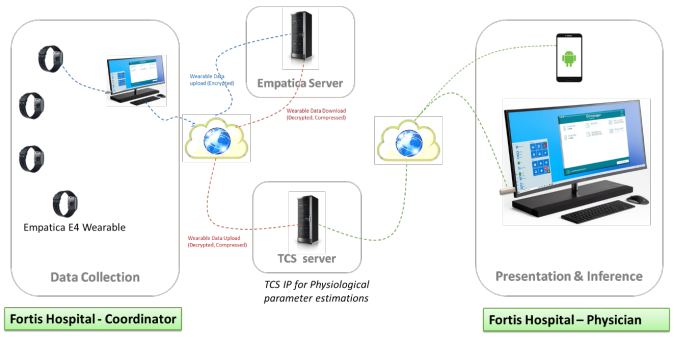


Fig. 1: System Deployment Architecture

as the photoplethysmography (PPG) and inertial measurement unit (IMU) that assist in computing metrics such as the heart rate and activity level. For our system, we used the Empatica E4 [8] primarily for its high level of data integrity and number of sensors available on the device. The list of sensors available on the device and the set of measures derived from them are highlighted in Table I.

The pilot study, that started in the mid week of June, 2020, involves 10 healthcare professionals (7 doctors, 2 nurses, 1 Technician) from Fortis hospital in India treating coronavirus patients. The device would be worn for 24hours (including during their 8-12 hour shifts) a day, except when there is some discomfort or during use of wash room/shower, with data synchronization and device charging/sanitization occurring prior to the start of their shift. At the time of onboarding, data such as blood pressure, body temperature (Axila), oxygen saturation (SpO2) and heart rate are captured using medical grade device to serve as a baseline measure. To maintain data privacy, the hospital staff handle mapping of device to participant and only a user ID stored in the database. The overall system deployment and data flow is captured in Figure 1.

Given the restriction of carrying a mobile device during their shift, and in particular within the covid ward, the E4 device operates in 'record' mode where data is stored locally on the device and later synchronized offline to the Empatica cloud. The Empatica portal allows us to download this synchronized data post which we upload (and process) on our EzyDoc analytics server. The EzyDoc portal also has an interface through which the hospital can view the different measures (see Table I) of their staff.

TABLE II: Cohort Statistics

Cohort	Start and End Date	Subject Count	Usable Duration (in hours)
C1	15^{th} June to 16^{th} July, 2020	3	645.71
C2	21^{st} July to 7^{th} August, 2020	4	678.68
C3	27^{th} Jan to 21^{st} Feb, 2021	8	1371.53
C4	19^{th} April to 5^{th} May, 2021	7	1058.66

III. Dataset Description

Data collection for all our subject was done using an off-the-shelf wrist sensing device called Empatica E4 [8]. The data avialable from this device are PPG at 64Hz, electrodermal activity (EDA) at 4Hz, 3-axis accelerometer at 32Hz, heart rate (HR) at 4Hz and temparature data at 1hz.

Data was collected as 4 cohorts (Table II), with different subjects in each cohort. The same device is not used for more than one user id in a given cohort. All the data were recorded on the device itself, which were pulled and synced to the database at the end of the day. Extra care was taken to ensure the watch is worn firmly on the wrist and should remain in contact with the skin for accurate PPG measurements. PPE kit was worn above the device.

For *Cohort 1* (C1), 20 participants signed up for the experiment. But due to non compliance (not wearing the watch or not following the protocol), limited data was found to be usable. Ultimately we ended up using data for 3 subjects w.r.t duration of data provided. Data collection for *Cohort 2* (C2) lasted for approximately 15 days. In this cohort, 6 subjects participated, but out of 6, data from only 4 subjects turned out to be useful. RT-PCR tests were conducted for the 4 subjects at the end of the cohort which resulted in 2 subjects being tested *COVID-19 positive* (COV+) and other 2 *COVID-19 negative* (COV-).

For *Cohort 3* (C3), comprising of 8 participants, data was collected in the same fashion and RT-PCR test was conducted regularly to evaluate the participants for any infection. *Cohort 4* (C4)comprised of 7 participants some of whom were infected while some were not based on RT-PCR test. The results of RT-PCR test for both C3 and C4 were not revealed to the researchers making it a double blind test. Once the data analysis was completed, the predictions for C3 and C4 were compared with the RT-PCR results for validation.

IV. Proposed Method/Methodology

Figure 2 shows a processing pipeline adopted. The collected data is first synced to the servers. Which is then used by the EzyDoc portal to estimate the vital parameters mentioned in Table I. Parallelly, it is used by the processing pipeline for further analysis.

The collected data is used to compute different parameters such as HR, RR etc. as mentioned in Table I. But these measures can only give a preliminary indicator of the well being of the participant. Onset of a disease is marked by certain prior symptons which leads to changes in the physiology of the person. In the case of COVID-19, whose

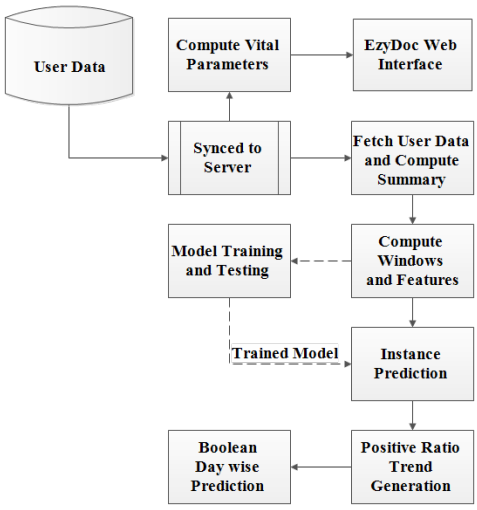


Fig. 2: Proposed processing pipeline

initial symptons are very similar to common flu, which is difficult to distinguish and pin point the actual cause. In this regard, knowing that the symptons have an effect on HR and oxygen saturation of an individual, its worth to study these dynamics during the infected period and be able to get some insight, so as to perform an early diagonosis. PPG has been used widely to study the above dynamics and is easily captured via wrist based wearable device such as Empatica E4.

From the above collected data, we use the PPG signal in a machine learning pipeline to build a model such that, the generated model can be used to classify whether new incoming data is from a healthy subject or not.

Model building data is taken as Cohort 2 data, which comprises of 2 subjects who were infected with COVID-19 virus during the process of data collection. The average duration of data obtained from each participant was approximately 100 hours across 15 days. Given the size of the data a window size of 5 mins was chosen for the analysis of the PPG signal. The raw data is pre-processed to remove any noisy or bad sections of the data. The conditioned signal is then divided into chunks of 5 minute windows. Each such window is then fed to an in-house tool called Feature discovery platform [11], which returns a set of 392 features comprising of time domain, frequency domain, wavelet transformed derivatives and hurst components.

A. Bin-wise approach to Machine learning model

A machine learning pipeline is then followed. All participants we screened for infection and declared as COV- before starting the experiment, confirming that all participants are healthy at the beginning. The data instances of each subject is divided into three *bins* (bin 1, 2 and 3) based on a set of days in chronological order. For example, as given in Table II a subject having 12 days of data is divided into 3 bins of 4 days each. Since some of the subjects tested positive for Covid-19 at the end of the cohort period, it assumed that they got infected towards the end of bin 2 or during bin 3. Hence we expect the physiological changes to be found during bin 3 for an infected subject. Another reson why we divided the data into 3 bins, was to be able to identify the inter bin differences and hence identify markers as early in the cycle as possible. To build the model we take bin 1 and bin 2 data for all the subjects of cohort 2. Instances were labeled as *COV+* or *COV-* as per the RT-PCR test results. Correspondingly, bin 3 data for all subjects is used as test instances. Finally the trained model is then tested on other cohort data as well.

B. Positve Instance Ratio Trend Generation

The proposed architecture aims to monitor a subject over a period of time and raise a flag if it detects some deviation from the normal. Here the deviation from the normal is treated as the classifier predicting the instances of the subject as positive class. But we cannot rely on just one instance prediction, which corresponds to only 5 mins of data from the user. Thus a continuous stream of the classifier predicting positive labels over a period of time is to be taken under consideration. For a given subject, we compute the *Positive Instance Ratio* for a duration, as given in equation 1, which gives us the ratio of the number of instances predicted as positive class with respect to total number of instances observed in say d hours of data.

$$PIR_t = \frac{\text{Number of positive prediction in } d \text{ hours}}{\text{Total number of instances in } d \text{ hours}} \quad (1)$$

at t^{th} timestep. PIR_t encompasses the data in the range of $[t-d, t]$ hours. From this, we generate trend line PIR_{avg} from PIR_t values by taking a moving average of window size $d = 24$ hours and sliding window size of $k = 6$ hours. Since data was not always continuous due to wash room / shower or other breaks, care was taken to consider only the available data points in a given time frame of $n = 24$ hours, for computation of moving average.

C. Boolean Daywise Prediction

From a system perspective, to have an automated system that can detect if a person is at risk, the following logic is used. If $mean(PIR_{avg}) \geq 0.6$ on i^{th} day, the system asserts a label of COV+ for that day otherwise COV-. Threshold of 0.6 is determined empirically, which is the mean of the PIR_{avg} for the COV+ subjects of C2.

V. Experiments and Results

Given the problem, our objective is to correctly detect subjects who are infected with the virus, i.e. positive class labels. Hence we look forward to obtain higher sensitivity of the learned model which means the model should **not** assert a infected subject as not infected, as it puts the subject at risk.

A. Learning a model for classification

To train a machine learning model, using cohort 2 data, we had data from 4 subjects. Each subject's data was divided into train and test set w.r.t 80-20 split criteria. 80% instances of each day is taken as train data (reffered as *C2-Train*) and

TABLE III: Model Training and Testing using 80% Cohort 2 data for training

TRAIN	TEST	#Feature	Accuracy	Sensitivity	Specificity
C2-Train	C2-Test	392	**0.8998**	0.7862	**0.9545**
C2-Train	C1-Test	392	0.7453	NA*	0.74539
C2-Train	C2-Test	15 MICe features	0.8481	0.7117	**0.9139**

*NA: not applicable as C1-Test had only one class data (COV-)

20% of each day data is taken as test data (reffered as *C2-Test*). This is done for all the subjects. Using this labeled train and test instances a machine learning model is trained. We used different machine learning models like support vector machines, k-nearest neighbour, logistic regression and random forest. Among them Random forest classifier with 100 trees gave us optimum results for different combination. As reported in Table III, using all the 392 features, we obtain an accuracy of 89.9% on *C2-Test* and 74.5% cohort 1 data (reffered as *C1-Test*).

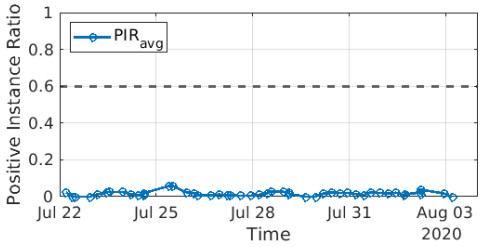
Feature selection is performed using Maximal Information Coefficient (MICe) [12], by which a subset of 15 features is selected. Using this set of 15 features and training a model, we obtained an accuracy of 84.8% on C2-Test data. It is observed that there is a drop in the metrics reported. The reduction in features was done to facilitate the system implementation. Computing 15 features is significantly faster and would enable predictions for a larger number of users. Since our system predicts the class on a daily basis , based on the entire day's data and a longitudinal threshold, marginal variation in accuracy will not affect the end result.

Figure 3 shows the positive ratio plots for two subjects under study. For C2-USR019 (Figure 3a), who was COV- throughout the study, has a PIR constantly below the threshold of 0.6. On the other hand for C2-USR021 (Figure 3b), who was infected during the study, has a PIR plot above 0.6 in the date range of July-August, 2020.

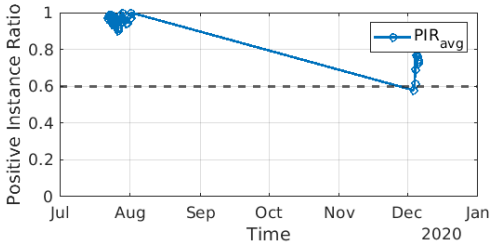
Further, to validate our model, we collected some more data from User 21, 24 during December 2020. Given the subjects have recovered completely, in Figure 3b, we see a drop in positive ratio values during the timeline of December 2020, further confirming the trained model is working correctly.

B. Unbiased testing of classifier

For Cohort 3 and Cohort 4, only the data for the entire cohort of 14 days were provided and not the ground truth. As these data sets were double blinded, they were used as test set data for our pipeline. The results of the corresponding RT-PCR test were not made available to the researchers apriori. Using our trained model so far and pipeline, the PIR plots, were obtained. Figure 4, provides the PIR plots for 2 users, where we can see the values remain below our threshold of 0.6. Hence as per the proposed pipeline we assert all the subjects of Cohort 3 as all COV-. Later the corresponding results of RT-PCR was made available where all the subjects were reported as COV-, matching

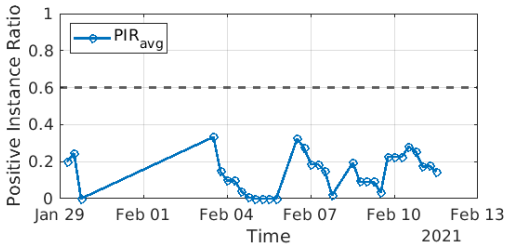


(a) Cohort 2 User 19, COV-

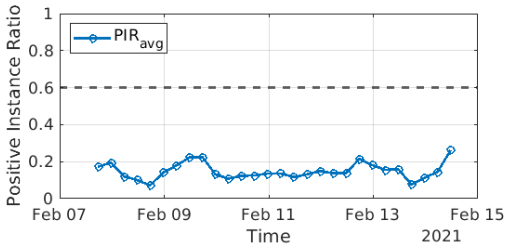


(b) Cohort 2 User 21, COV+

Fig. 3: Cohort 2 Positive Instance Ratio plots



(a) Cohort 3 User 7, COV-



(b) Cohort 3 User 11, COV-

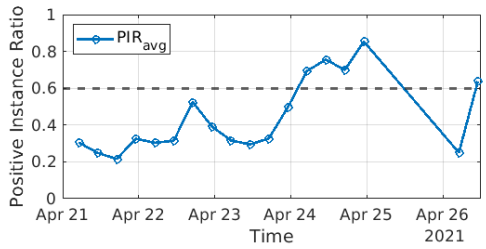
Fig. 4: Cohort 3 Positive Instance Ratio plots

our predictions. Being frontline workers, all the cohort 3 participants were vaccinated and hence did not contract the infection.

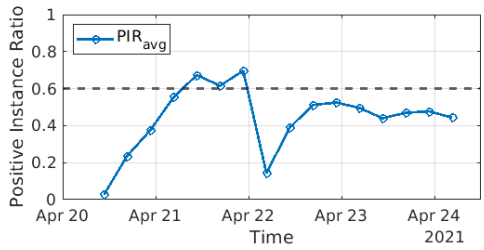
Similarly for Cohort 4, the prediction made by our pipeline are listed in Table IV and the corresponding plots for user 1 and user 5 are given below (Figure 5). Out of the 7 subjects, our system predicted 4 as COV+ and 3 as COV-. When the RT-PCR results were revealed, 4 subjects turned out to be COV+ and the remaining 3 COV-, matching our predictions.

VI. Conclusion

This paper proposes a pipeline to detect abnormalities from a normal behaviour, based on trained classifier and smoothened version of longitudinal prediction over a period



(a) Cohort 4 User 1, COV+



(b) Cohort 4 User 5, COV+

Fig. 5: Cohort 4 Positive Instance Ratio plots

TABLE IV: Cohort 4 Evaluation

User ID	Ground Truth	Prediction
C4-USR001	COV+	COV+
C4-USR002	COV+	COV+
C4-USR003	COV+	COV+
C4-USR004	COV-	COV-
C4-USR005	COV+	COV+
C4-USR006	COV-	COV-
C4-USR009	COV-	COV-

of time. In this paper we demonstrate the same with a use case of current COVID-19 pandemic, where data from several front line workers from a reputed hospital in India is collected, analysed to build a machine learning model which is then used further to obtain prediction on a longitudinal basis to assert whether a subject is at risk. This pipeline is not limited to COVID 19 pandemic but can be re-used to obtain a generic framework for such other use cases. Proposed systems is primarily for HCP as they are at the highest risk due to continuous exposure. The idea is to boost their confidence by providing a system for early warning. Patients have already turned positive and hence such monitoring is not useful. Furhter the system is not limited to only frontline workers but can easily reused for general consumers for longitudinal monitoring.

We would further like to make it clear, as per above use case, our objective is not repalce RT-PCR tests, by longitudinal monitoring via a wearable device, instead we want develop a system that can acts as an assitive aid for the subject to monitor the vitals as well as raise alerts for the subject when the pipeline forsees a risk. Our key contribution is a system that can predict the onset of the infection trend a few days in advance of the actual clinical deterioration/classical symptoms appear. As we know, in asymptomatic patients, RT-PCR becomes positive in 5 to 7 days of contracting the virus. The alerts can be used to initiate a standard physical examination and RT-PCR test to confirm the findings leading to early interventions and thereby avoiding clinical complications of the disease such as cytokine storm.

In our study, we have proposed a boolean model and architecture for implementation in an automated fashion, along with how we have overcome the technical and non-technical challenges. Although our results are very good, the number of subjects available for training and testing are very limited. Hence the results have to be further validated on larger datasets. Currently the final verdict of a subject being Covid positive is given based on mannual observation of the positive ratio plots which needs to be automated for future needs.

REFERENCES

[1] Samiksha Goel. The doctor-population ratio in india is 1:1456 against who recommendation, January 2020.

[2] Amit Kumar Das, Sayantani Ghosh, Samiruddin Thunder, Rohit Dutta, Sachin Agarwal, and Amlan Chakrabarti. Automatic covid-19 detection from x-ray images using ensemble learning with convolutional neural network. *Pattern Analysis and Applications*, pages 1–14, 2021.

[3] Dandi Yang, Cristhian Martinez, Lara Visuña, Hardev Khandhar, Chintan Bhatt, and Jesus Carretero. Detection and analysis of covid-19 in medical images using deep learning techniques. *Scientific Reports*, 11(1):1–13, 2021.

[4] Afshin Shoeibi, Marjane Khodatars, Roohallah Alizadehsani, Navid Ghassemi, Mahboobeh Jafari, Parisa Moridian, Ali Khadem, Delaram Sadeghi, Sadiq Hussain, Assef Zare, et al. Automated detection and forecasting of covid-19 using deep learning techniques: A review. *arXiv preprint arXiv:2007.10785*, 2020.

[5] Jennifer M Radin, Nathan E Wineinger, Eric J Topol, and Steven R Steinhubl. Harnessing wearable device data to improve state-level real-time surveillance of influenza-like illness in the usa: a population-based study. *The Lancet Digital Health*, 2(2):e85 – e93, 2020.

[6] H Ceren Ates and et.al. Wearable devices for the detection of covid-19. *Nature Electronics*, 4(1):13–14, 2021.

[7] Gireesh K Bogu and Michael P Snyder. Deep learning-based detection of covid-19 using wearables data. *MedRxiv*, 2021.

[8] Empatica, https://www.empatica.com/embrace2/.

[9] Vivek Chandel and et.al. C2p: An unobtrusive smartwatch-based platform for automatic background monitoring of fatigue. In *Proceedings of the First International Workshop on Human-Centered Sensing, Networking, and Systems*, HumanSys'17. Association for Computing Machinery, 2017.

[10] Vivek Chandel and et.al. Actrak-unobtrusive activity detection and step counting using smartphones. In *MobiQuitous*. Springer International Publishing, 2014.

[11] Arijit Ukil and et.al. Automodeling: Integrated approach for automated model generation by ensemble selection of feature subset and classifier. In *2018 International Joint Conference on Neural Networks (IJCNN)*, pages 1–8. IEEE, 2018.

[12] David N Reshef and et.al. Detecting novel associations in large data sets. *science*, 334(6062):1518–1524, 2011.