SAMHAR COVID-19 HACKATHON

Problem Statements

Title	Drug Discovery and Genome Sequencing
Problem Statement	The genome of novel Coronavirus has been sequenced by various countries, and reference of same been built having approximate size of 30kb. India became the fifth country in the world to sequence the genome of the novel Coronavirus, or COVID19, and share its data with the international community. Two sequence from India has been submitted to the NCBI database. Genome data is essential to build tests, find drugs and vaccines. As virus tend to mutate at a high rate and hence can become more pathogenic. Hence understanding of changes in virus genome becomes essential. The study on changes in virus genome can also guide in determining the source of infection and thereby will help in spread of Virus. The study of genome will also help in finding unique regions in the virus genome and which can target for drug development. Drug Development typically takes at least decade to move from idea to market. Also, crystal structures for some of the proteins for coronavirus has been solved. Hence Drug repurposing studies can be carried out using already FDA approved ligand database. We can substantially accelerate this process using AI and make it much cheaper, faster, and more likely to succeed. The objective of this problem statement is to find novel ideas to identify potential drugs and drug targets for the treatment of COVID-2019. Expected Outcome: • AI based tool to identify novel drug targets based on genome data • To develop AI based tool for faster screening of available ligand molecules to predict probable lead molecules (Drug Repurposing).
Sample Dataset for Reference	• https://www.ncbi.nlm.nih.gov/genbank/sars-cov-2-seqs/ • https://www.gisaid.org/epiflu-applications/next-hcov-19-app/ http://www.rcsb.org/ **Indian Submissions**: Complete Genome • https://www.ncbi.nlm.nih.gov/nuccore/MT012098 • https://www.ncbi.nlm.nih.gov/nuccore/MT050493 **Indian Submissions**: Partial CDS • https://www.ncbi.nlm.nih.gov/nuccore/MT163715 • https://www.ncbi.nlm.nih.gov/nuccore/MT163714 These are links to sample dataset for reference purpose only. Participants are free to use these or other datasets after studying respective licenses.
References of Research papers/ Technical	• https://community.topcoder.com/longcontest/?module=ViewProblemStatement&rd=16555&compid=49304

documents	• https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=6451097 • https://spectrum.ieee.org/the-human-os/artificial-intelligence/medical-ai/companies-ai-coronavirus • https://www.ipd.uw.edu/2020/02/rosettas-role-in-fighting-coronavirus/ • https://www.nature.com/articles/s41586-019-1923-7.epdf • https://github.com/deepmind/deepmind-research/tree/master/alphafold_casp13

Title	**Diagnosing COVID-19 using AI-based medical image analyses**
Problem Statement	Radiologists have watched the coronavirus disease 2019 (COVID-19) pandemic unfold. Radiology literature suggests a pivotal role for CT/X-RAY as CT/X-Ray findings in Covid-19 patient has pneumonia, and hence medical imaging has high sensitivity for diagnosis of COVID-19. The objective of this statement is to figure out usage of AI in medical imaging modalities for fast and accurate diagnosis of COVID-19 **Expected Outcome**: Diagnosing or detecting patients infected with COVID-19 using chest radiography images
Sample Dataset for Reference	• https://www.kaggle.com/paultimothymooney/chest-xray-pneumonia • https://github.com/lindawangg/COVID-Net • https://github.com/ieee8023/covid-chestxray-dataset • https://www.rsna.org/en/education/ai-resources-and-training/ai-image-challenge/RSNA-Pneumonia-Detection-Challenge-2018 These are links to sample dataset for reference purpose only. Participants are free to use these or other datasets after studying respective licenses
References of Research papers/ Technical documents	• https://www.quantib.com/blog/diagnosing-covid-19-using-ai-based-medical-image-analyses • https://pubs.rsna.org/doi/10.1148/radiol.2020200905 • https://arxiv.org/pdf/2003.09871.pdf

Title	**Public Health Surveillance and Monitoring**
Problem Statement	In order to enforce social distancing, it is important to monitor suspicious scenarios like crowd gathering or tracking etc. Usage of AI (Convolution Neural Network) has surpassed traditional Computer Vision techniques and can help automate this pattern recognition. The objective of this problem statement is to find different pattern on input sources like camera, satellite, social platforms which can help identify suspicious activities to stop or track covid-19 spread.
Sample Dataset for Reference	**Crowd Counting/Gathering Database**: http://cs-chan.com/project4.htm https://riemenschneider.hayko.at/vision/dataset/index.php?filter=+crowd https://github.com/gjy3035/Awesome-Crowd-Counting http://gamma.cs.unc.edu/LCrowdV/#dataset http://gamma.cs.unc.edu/GAIT/##EWalk http://gamma.cs.unc.edu/RCrowdT/ **People Detection and Tracking**:

	http://homepages.inf.ed.ac.uk/rbf/CVonline/Imagedbase.htm#people **Flu Recognition**: https://web.bii.a-star.edu.sg/~chengli/FluRecognition.htm These are links to sample dataset for reference purpose only. Participants are free to use these or other datasets after studying respective licenses
References of Research papers/ Technical documents	• https://nanonets.com/blog/crowd-counting-review/ • https://www.analyticsvidhya.com/blog/2019/02/building-crowd-counting-model-python/

Title	**Pandemic Forecasting**
Problem Statement	The number of new cases is increasing day by day around the world. The participants can use the dataset that has information from the states and union territories of India daily and in some case at hourly intervals. In this challenge, you will be predicting the cumulative number of confirmed COVID19 cases in various locations across the world, as well as the number of resulting fatalities, for future dates. Also, this can be combined with social dataset like of Twitter where the first tweet goes back to January 22, 2020. Other parameters of importance could be the impact of weather and India Population Demographics.
Sample Dataset for Reference	• https://www.covid19india.org/ • https://coronavirus.jhu.edu/map.html https://github.com/echen102/COVID-19-TweetIDs • https://archive.org/details/twitterstream • https://www.meteoblue.com/en/weather/archive/export/india_el-salvador_3585481 • https://www.indiastat.com/demographics-data/7/stats.aspx • http://www.mospi.gov.in/download-tables-data • https://www.bing.com/covid • https://www.acaps.org/covid19-government-measures-dataset These are links to sample dataset for reference purpose only. Participants are free to use these or other datasets after studying respective licenses
References of Research papers/ Technical documents	• https://healthcare-in-europe.com/en/news/predicting-the-future-of-the-covid-19-pandemic-with-data.html

Title	Data Mining
Problem Statement	With Covid-19 becoming a pandemic the research on same has also accelerated. There are more than 44000 research articles already published and growing. It is important to implement functionalities like Data mining that can help the medical community to develop answers to high priority scientific questions. AI can help mine through news reports and online content from around the world, helping experts recognize anomalies even before it reaches epidemic proportions. The objective is to analyze the existing literature for analysis including question, answering and summarization which helps in getting better and faster insights.
Sample Dataset for Reference	• https://www.ncbi.nlm.nih.gov/pmc/?term=covid-19 • https://www.ncbi.nlm.nih.gov/pubmed/?term=covid-19 • https://www.cdc.gov/library/researchguides/2019novelcoronavirus/researcharticles.html These are links to sample dataset for reference purpose only. Participants are free to use these or other datasets after studying respective licenses
References of Research papers/ Technical documents	• https://www.cdc.gov/library/researchguides/2019novelcoronavirus/databasesjournals.html

Title	Other/Open House Challenge
Problem Statement	The participants may bring their own problem statement along with datasets and apply with solutions with the following priorities (but not limited to): • To achieve better (More Accurate) and Faster Turnaround time for Testing and results. • To identify, track and forecast outbreaks. • To develop the Drug Re-purposing Simulations leading to New Drug Discovery. • To develop the Simulations/Solutions using Traditional (AYUSH) Medicines. • To create Chat Bot with multiple regional language supporting Conversational AI Techniques. • To develop AI based identification of non-complying or infected individuals. • To create, Train and Deploy Robots to sterilize, Deliver Food and Supplies. • To deploy Drones to monitor sites, announce policy measures and to deliver medical supplies in the infected regions/areas; To enable automation of COVID 19 Health care Claims for

	Ayushman Bharat Program.