

# GYANENDRA DAS

## Kaggle Competition Expert

+91-9178158633 ◇ gyanendra.19je0343@am.iitism.ac.in

## ACHIEVEMENTS

---

1. **Kaggle Competitions Expert** - Highest Rank 1513 out of 157061 active Kagglers
2. **Samsung Innovation Award 2020** Finished 1st rank among 1k+ teams
3. **Flipkart Grid 2.0** - National Finalist

## KAGGLE

---

**Ranked Top 1 % Globally in Kaggle Competitions** ([www.kaggle.com/gyanendradas](http://www.kaggle.com/gyanendradas))

Competition	Medal	Rank(percent)	Rank
<b>SIIM-ISIC Melanoma Classification</b>	GOLD	Top 1%	6/3308
<b>Cassava Leaf Disease Classification</b>	SILVER	Top 2%	42/3947
<b>Tweet Sentiment Extraction</b>	SILVER	Top 5%	128/2227
<b>University of Liverpool - Ion Switching</b>	BRONZE	Top 9%	218/2618
<b>Riiid Answer Correctness Prediction</b>	BRONZE	Top 9%	286/3395

## EDUCATION

---

**Integrated Master of Technology in Mathematics and Computing**  
IIT(ISM), Dhanbad. CGPA :7.76/10

*2018-Present*

## WORK EXPERIENCE

---

### **Willow.AI**

*Deep Learning Research Intern*

August - December 2020

*Guragaon, India*

- Elliot waveform recognition in stock market data. Our goal is to find and identify important Elliott volatility charts, and use this function to predict the target range and the probability of success for the next trader. We propose an end-to-end trainable Deep Learning Neural Network and Recurrent Neural Network. Extensive experiments show that the proposed method consistently outperforms state-of-the-art methods. Our method achieves 94.37

### **Travel Buddy**

*Machine Learning Intern*

May - June 2020

*Guragaon, India*

- Provide two real time working model. One of them is Not safe for Work classification for both image and video. Second one is Toxic text Classification supporting multi languages. Challenge for Text is it has very imbalance data-set Overcome this issue and Transformer Model is used and got ROC accuracy of 94.3. For image Efficient Net is used and For video Efficient Net with RNN is used.

### **B3 Digital Solution**

*Data Science Intern*

October 2019 - March 2020

*NCR Delhi, India*

- Analysis of business data and maximize profit. Also work on a project to Detect Breast Cancerous Tissue with image processing and it's location and shape and provide one real time working model to detect anonymous chat.

## PROJECTS

---

### **Invo AI-An AI based E-Invoicing Solution**

Nov 2020 - Nov 2020

- We made a model which can extract all information in a scanned invoice to a XML file. We use various image preprocessing Task and Deep learning model. It is very time efficient and computer resource efficient. The highest accuracy we achieved is 87.6 (precision) which out perform Google cloud vision 80 (precision). This project leads us to win Samsung innovation Award 2020 and Flipkart GriD 2.0.

### **Brightfield to Fluorescent Microscopy Data**

Oct 2020 - Ongoing

- Using GAN and Sequential Model we are targeting to convert Brightfield Image of Microscopy data to fluorescent so that we can beforehand which area we have to swell up when applying green dye. We achieved 94.37 accuracy till now. We are working on this.

### **SIIM ISIC Melanoma Classification**

June 2020 - August 2020

- It was a Kaggle Competition and in Private leader board I stand 6 out of 3315 participants and earned a gold model. The Goal in this Competition to create a model which can classify the melanoma cell correctly. Here The metric used was AUC and our score was 0.9470.

### **University of Liverpool - Ion Switching**

March 2020 - May 2020

- It was a Kaggle Competition and in Private leader board I stand 218 out of 2618 participants and earned a silver model. In this competition, we have to use ion channel data to better model automatic identification methods. If successful, we would be able to detect individual ion channel events in noisy raw signals. Here The metric used was Micro F1score and my score was 0.94397.

### **Corona-Drug-Discovery**

May 2020 - June 2020

- This is a solution for the Possible Drugs for Covid-19 . Binding scores of leading existing drugs (HIV inhibitors) are around -10 to -11 and around -13 for the drug Remdesivir which recently entered clinical testing. More negative the binding score is, better the drug is. The goal is to create a novel small molecule which can bind with the coronavirus, using deep learning techniques for molecule generation. I was able to create several small molecule candidates which achieved binding scores approaching -18.

### **Protein-Structure-Prediction**

May 2020 - July 2020

- Rapid prototyping of ideas I have regarding protein structure prediction, some of them building on top of the idea of molecular dynamics simulations, others from the ground up. It may very well be that none of these ideas ever even achieve to fold a simple protein but iterating over different ideas and implementing the ba...

### **Duplicacy - Quora Question Duplicate or Not**

Feb 2020 - March 2020

- The goal is to create a Deep learning model that can learn from the existed quora question and predict the input question is similar to our database questions or not. I distinguish three kind of features embedding features, classical text mining features and structural features. I worked on two main architectures for NNets : Siamese and Attention Neural Networks

## TECHNICAL STRENGTHS

---

### **Computer Languages**

C/C++, Python, R

### **Libraries**

OpenCV, Selenium, Pandas, Numpy,

### **Version Control**

Git, Github, VS-Code

### **Machine Learning**

Tensorflow, Keras, PyTorch, Reinforcement learning

### **Data Visualisation**

SQL, Tableau, Matplotlib, Seaborn, Scipy

### **Website (coded)**

HTML5 and CSS