Rainfall prediction for the state of Gujarat using deep learning technique

Rushikesh Nalla*, Urmil Kadakia†, Ranendu Ghosh‡
Dhirubhai Ambani Institute of Information and Communication Technology (DA-IICT), Gandhinagar, Gujarat
Email: *rushikeshnalla9@gmail.com, †urmil.kadakia@gmail.com, ‡ranendu_ghosh@daiict.ac.in,

***Abstract*—Prediction of rainfall which varies both spatially and temporally is extremely challenging. Infrared and visible spectral data from satellites have been extensively used for rainfall prediction. In this study, two deep learning methods MLP and LSTM are discussed at length for predicting precipitation at a fine spatial (10km×10km) and temporal (hourly) resolution for the state of Gujarat. These methods are applied by using the multispectral (VIS, SWIR, MIR, WV, TIR1, TIR2) channel data such as cloud top temperature and radiance values of the INSAT-3D satellite (ISRO) as features for the model. Textural features of satellite images are incorporated by considering mean and standard deviation of each pixel's neighbourhood. Rainfall also heavily depends on the elevation and vegetation of earth's surface so we have used SRTM DEM and AWIFS NDVI respectively. Measurements of actual rainfall are obtained from AWS (point source stations) and TRMM (10km×10km resolution). First dataset contains only TIR1 band temperature and AWS rainfall data for training but the second dataset includes multispectral channel data and TRMM rainfall data which brought about great improvement in results. For each data-set, a comparison between MLP and LSTM models is discussed here. We were able to classify the rainfall into nil (0mm), low (<2mm), medium (>=2mm and <5mm) and high (>=5 mm) with a high accuracy. Metrics like accuracy, precision, recall and fscore have been computed to get better insights about the dataset and its corresponding outcome. Our results show that LSTM performs significantly better than MLP for any given balanced class data-sets.**

***Keywords*—INSAT-3D, Multispectral channel, Rainfall, Normalized Difference Vegetation Index (NDVI), Shuttle Radar Topography Mission (SRTM), Solar Zenith Angle (SZA), Multi Layer Perceptron (MLP), Long Short Term Memory Module (LSTM), Automatic Weather Station(AWS), Tropical Rainfall Measurement Mission (TRMM)**

I. Introduction

Prediction of rainfall in terms of its amount is extremely challenging. Rainfall varies both spatially and temporally and it is useful in many areas ranging from flood and storm forecasting to climate modeling. There has been a significant improvement in the prediction of precipitation in the last two decades with the advancements in the field of satellite, radar and other observation techniques, algorithms and processing power. Still there is tremendous scope to improve upon it and to predict rainfall in finer spatial and temporal resolution.

The Infrared(IR 11μm) data from geosynchronous satellite is related to the cloud-top temperature. Generally it is assumed that intense rainfall is related to cold cloud-tops bright temperatures so it has greater probability of receiving rainfall. When we consider data with low spatial and temporal resolution it is easier to predict rainfall as compared to high spatial and temporal resolution data. This is due to the fact that rainfall is highly variable in the latter case. The former results in error cancellation and hence better results [1]. Rainfall and cloud top temperature are not directly (or inversely) related but has a complex relationship. The high-altitude cirrus cloud-tops have very low temperature so IR-based algorithm predicts them to have high rainfall always but this is not true. The passive microwave(PMW) gives better rainfall prediction compared to the previous method as it captures hydro-meteor related parameters information more accurately. They are carried via the low earth satellite so they have low resolution in the spatial and temporal dimension with low sampling frequency [2]. As both of the above methods results are seen at larger time instances (4-6 hours) they can not be used for applications such as flash flood prediction that happens in a time frame of an hour or two.

Some algorithms such as CMORPH (Climate Prediction Center morphing method) and MIRA (microwave/infrared rainfall algorithm) give good results by combining PMW data with the IR data. [3–6] but these methods suffer from similar problems related to that of PMW based methods. This happens because PMW satellite takes around 3 hours to scan a large percentage of the earth surface. The algorithm can not provide any results for the in between scenario. While the high spatial resolution of VIS(Visual) and IR(Infrared) coupled with the higher sampling frequency of Geo-satellites can capture the temporal variability that is useful for many applications. [7] These methods have their own set of issues. The visible data is not available throughout the day and the mapping of IR brightness temperature(Tb_{12}) to precipitation probability using Arkin-Meisner method [1] is not accurate in tropical region due to the non-convective high altitude cirrus clouds.

In the last two decades the quality of GEO-satellite observations has increased significantly. One way to tackle the rainfall detection and prediction problem is to use multispectral satellite data. For example, INSAT-3D (Indian National Satellite System) has 25 spectral bands ranging from 0.52μm to 14.71μm and currently scans earth every 30 minutes with a pixel resolution of 2km×2km. There have been multiple efforts to predict rainfall using multiple channels. A few studies have used single visual and single infrared channel[8]. In his paper, Toshiyuki Kurino [9] has argued that difference between brightness temperature of channel 11μm and 12μm (Tb_{11} - Tb_{12}) is useful for identifying thin cirrus clouds (no precipitation) and the difference between of brightness temperature of channel 11μm and 6.7μm (Tb_{11} - $Tb_{6.7}$) is useful for identifying deep convective clouds (heavy precipitation). GOES (Geostationary Operational Environmental Satellite system) Multispectral Rainfall Algorithm (GMSRA) [10] uses combined information from 5 channels 0.65μm, 3.9μm, 6.7μm, 11μm, 12μm coupled with pre-calibrated probability of rain derived from clouds top brightness temperature groups to predict rainfall. Some other techniques such as Self-Calibrating Multivariate Precipitation Retrieval (SCaMPR) [11] uses linear regression, Rainfall prediction from Remotely Sensed information using Artificial Neural Networks Multispectral Analysis (PERSIANN-MSA) [12] uses artificial neural network based self-organizing map(ANN-SOFM) for predicting rainfall.

II. Dataset Description

INSAT-3D is a multipurpose geosynchronous satellite launched by ISRO with two main payloads of IMAGER ans SOUNDER. The INSAT-3D provides VHRR (Very High Resolution Radiometer) data of multiple wavelength on half-hourly basis. MOSDAC has provided us with INSAT-3D data for each hour of each day for rainy months i.e. June, July, August, September from 2014 to 2017. Of the 25 spectral channels 5 channels are used in the present study, with wavelengths 0.52μm-0.72μm VIS (Visible) FIG. 1, 1.55μm-1.70μm SWIR (Short Wave Infrared)FIG. 2, 6.50μm-7.00μm WV (Water Vapor)FIG. 4, 10.2μm-11.2μm TIR-1 (Thermal Infrared)FIG. 5, 11.5μm-12.5μm TIR-2 (Thermal Infrared)FIG. 6. The temporal resolution is 1 hour and spatial resolution 2 km for all spectral channels.[13]

TIR1 spectral channel's radiant energy is proportional to fourth power of cloud top temperature. Using cloud top temperature we can find out brightness temperature that is indirectly related to the rainfall. Clouds having cloud top temperature less than 235K has high rainfall probability [13]. TIR2 in combination with TIR1 provides information related to cloud thickness. Thin clouds can not hold as much water as thicker ones [14]. The atmosphere has water in gaseous form. This has both vibrational and rotational transitions that gives rise to vibration-rotation spectrum. These spectral lines have the same frequency and energy as that of microwave and water-vapor spectrum. More the amount of gaseous water in the region the stronger is the absorption of WV and microwave radiation. [15] The clouds that are optically thick in the visible band are good candidates for rainfall. Emitted or reflected radiation originating from bodies below of clouds means that the clouds are either broken or semi-transparent which implies that the clouds are thin and hence less rainfall [16].

We have not used the other 20 channels as 13 of them are not related to the rainfall and rest of the channel have many entries that are missing or have been filled with default values. Some studies support the importance of channel MIR (3.8μm to 4.0μm) in the prediction of rainfall[14]. Despite this fact we have not used it in our study. The 3.8μm-4.0μm spectrum during the day contains thermal emission and solar reflection. Now to eliminate the effect of the solar reflection we have to identify and separate it and perform the correction related to the solar zenith angle(SZA). There has been several studies in the past on this topic but most of them have some assumption or simplification to ease the calculation. Further, we also have to do the corrections related to the thermal emission. Thus, we are not using MIR channel in the present study to prevent any misinterpretation.

The study region covers longitude 68°W-75°E and latitude 25°N-20°W (State of Gujarat, India). This region suits our purpose as it is close to the equator where most of the precipitation events occur. Close to the equator the solar zenith angle becomes more significant compared to higher latitudes.

As wavelength increases, the effect of solar zenith angle decreases gradually. As VIS (0.52μm-0.72μm) and SWIR (1.55μm-1.70μm) have low wavelength, they are highly effected by the solar zenith angle (SZA) so the corrections are necessary. The method of calculating Solar Zenith Angle is shown in FIG. 9. There are many methods for doing this correction [17, 18]. According to the previous studies, a reasonable correction technique is to multiply the observed value by its associated $(\cos SZA)^{-1}$. Because of the inconsistency associated with SZA $> 60°$, normalization only applies to the values with the SZA $< 60°$.

SRTM (Shuttle Radar Topography Mission) provides DEM (Digital Elevation Model) for the entire globe. The DEM has 1km$\times$1km resolution and the vertical error for DEM is less than 16m[19]. The SRTM map of Gujarat is shown in FIG. 8. The NDVI (Normalized Difference Vegetation Index) is a numerical indicator that shows whether targeted area contains green vegetation or not, The NDVI map of Gujarat is shown in FIG. 7. Oceansat-2 Ocean Color Monitor (OCM2) Global Area Coverage (GAC) sensor is used to generate NDVI products for a 15 day period. As the OCM2 is a swath imaging system it is necessary to apply corrections related to solar zenith angle, cloud masking and surface reflectance. Then NDVI is calculated from the atmospheric reflectance of the NIR and red band. Finally, the NDVI images are stacked to create a 15 day period image with the resolution of 1km$\times$1km. [20]

Here we have used ground based rainfall observation stations (AWS - Automatic Weather Station) in the first part of the study. This data is also provided by the MOSDAC website. The temporal resolution of the data is 1 hour but it has no specific spatial resolution as the station is a discrete point of the ground. The distances between stations are not uniform. The spatial resolution of IR is 2km$\times$2km. Here we have taken the grid points that includes a station in it and assumes that the grid's rainfall is same as station's rainfall measurement. In second case, due the limited amount of data provided by the ground stations we use TRMM PR data for training and validation. In the second case the spatial resolution is 0.1 degree(10km$\times$10km) as the resolution of TRMM is 0.1 degree.

Dataset	**Band Wavelength (μm)**	**Features per channel**	**Other Features**	**Input Dimension**
1	10.8	5	3	8
2	0.65 + 1.6 + 6.2 10.8 + 12	5	2	27

TABLE I: Features used for each data-set

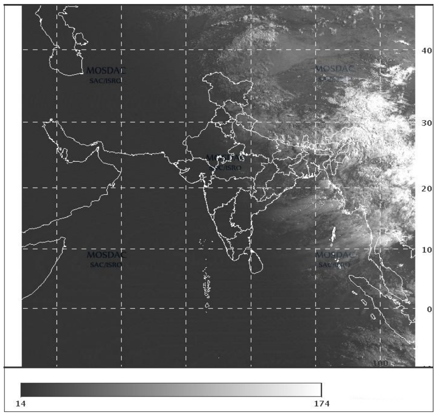

Fig. 1: INSAT-3D (IMAGER), Wavelength = 0.65μm (VIS), Date = 06\07\2017, Time = 10:00 GMT

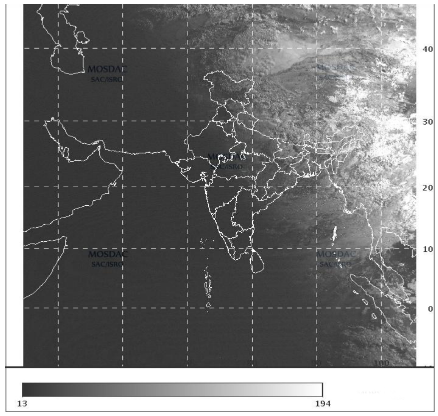

Fig. 2: INSAT-3D (IMAGER), Wavelength = 1.625μm (SWIR), Date = 06\07\2017, Time = 10:00 GMT

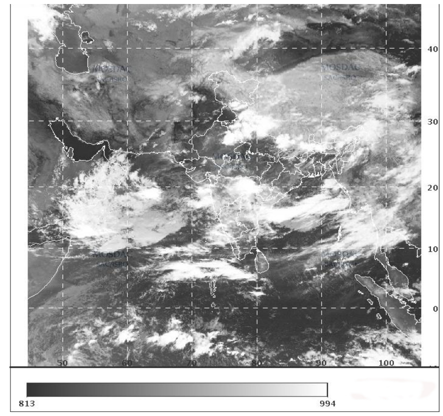

Fig. 3: INSAT-3D (IMAGER), Wavelength = 3.9μm (MIR), Date = 06\07\2017, Time = 10:00 GMT

[13]

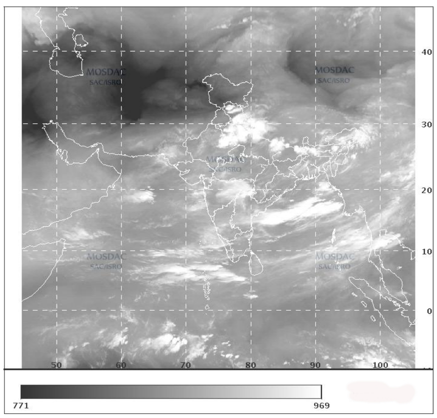

Fig. 4: INSAT-3D (IMAGER), Wavelength = 6.8μm (WV), Date = 06\07\2017, Time = 10:00 GMT

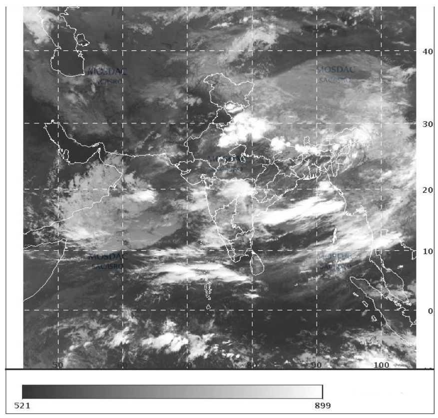

Fig. 5: INSAT-3D (IMAGER), Wavelength = 10.8μm (TIR1), Date = 06\07\2017, Time = 10:00 GMT

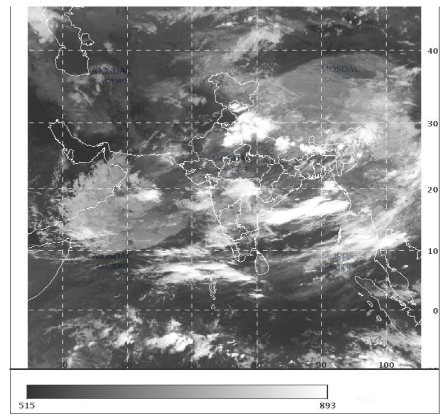

Fig. 6: INSAT-3D (IMAGER), Wavelength = 12μm (TIR2), Date = 06\07\2017, Time = 10:00 GMT

[13]

III. Data Preprocessing

A. Dataset 1

- Infrared Satellite Images (ASIA) were taken from the MOSDAC data repository which was nearly 500GB in size.
- Extracting values corresponding to the state of Gujarat.
- Creating an intermediate dataset containing latitude, longitude and the corresponding cloud top brightness temperature (TIR1 band 10.8μm).
- Computing the brightness temperature mean and standard deviation of 3$\times$3 neighbourhood.
- Computing the brightness temperature mean and standard deviation of 5$\times$5 neighbourhood.
- Up-scaling SRTM Digital Elevation data (1km$\times$1km) resolution data to match with IR Satellite Data (2km$\times$2km) resolution.
- Up-scaling AWIFS Normalized Difference Vegetation Index (1km$\times$1km) resolution data to match with IR Satellite Data (2km$\times$2km) resolution.
- Cleaning AWS (Automatic Weather Station) Rainfall data.
  - Deleting garbage values and filling inconsistent values with most suitable average.
  - Places with large discrepancies were removed completely.
  - Putting various checkpoints and verifying manually by comparing it with IMD (Indian Meteorological Department) rainfall data.
- Mapping intermediate dataset with up-scaled values of SRTM DEM, AWIFS NDVI, neighbourhood features and AWS Rainfall.

B. Dataset 2

- Extracting TIR1 (10.8μm), TIR2 (11.9μm), MIR (3.9μm) and WV (6.8μm) Temperature values.
- Computing the brightness temperature mean and standard deviation of 3$\times$3 neighbourhood for above bands.
- Computing the brightness temperature mean and standard deviation of 5$\times$5 neighbourhood for above bands.

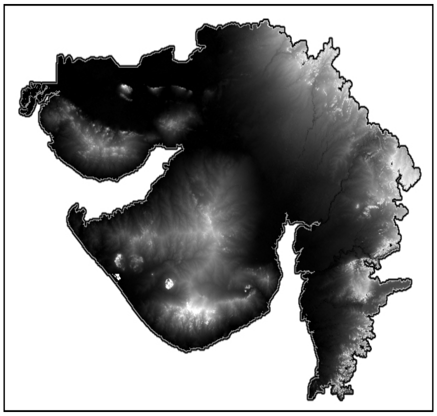

Fig. 7: SRTM Digital Elevation Model data. White = max height, Black = min height

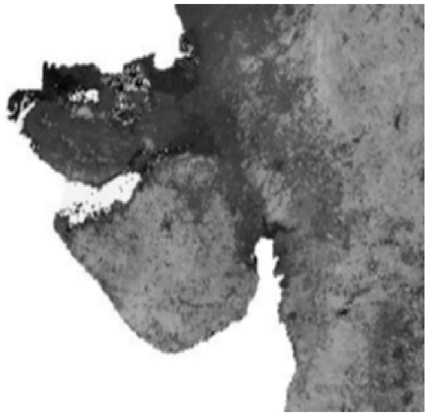

Fig. 8: NDVI (Filtered Normalized Difference Vegetation Index), Usage Range = 0-200 white = outside country boundary [13]

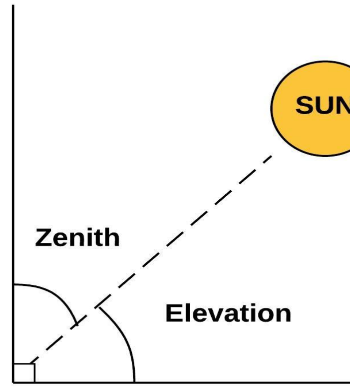

Fig. 9: Solar Zenith Angle and Elevation

- Extracting SWIR (1.6μm) and WV (6.8μm) Radiance values.
- Extracting VIS (0.6μm) Albedo values.
- Computing the radiance mean and standard deviation of 3$\times$3 neighbourhood for SWIR, WV and VIS.
- Computing the radiance mean and standard deviation of 5$\times$5 neighbourhood for SWIR, WV and VIS.
- Each of the above mentioned features have (2km$\times$2km) resolution.
- Up-scaling SRTM Digital Elevation data (1km$\times$1km) and AWIFS Normalized Difference Vegetation Index (1km$\times$1km) resolution data to match with TRMM (10km$\times$10km) resolution data.
- Extracting TRMM Rainfall data which has (10km$\times$10km) resolution.
- Up-scaling all the features to map with TRMM Rainfall data.

IV. Problems in AWS Rainfall Data

- Rainfall data is cumulative and the reset points are at random.
- Appearance of 1023 due to two reasons:
  - It is the highest value (10 bit number).
  - Missing data is sometimes given the value 1023.
- Random length of increasing and decreasing numbers.
- Garbage values like 9999 appear randomly.
- Nearly $(1/3)^{rd}$ data is missing.

V. Implementation

We started by developing some linear classifier models but didn't get satisfactory results as estimating rainfall is a complex problem. So, we moved on to non-linear classifiers which are capable of solving intricate issues. In this study, 80% of total data was allocated for training and rest (20%) was used for validation.

We have used NVIDIA GEFORCE GTX 1080 graphics card which has 2560 cores, 8GB GDDR5X RAM and memory speed of 10Gbps for implementation. The processor used is Intel(R) Core(TM) i7-8700K CPU @ 3.70GHz with 16GB RAM.

A. Multi-layer Perceptron (MLP)

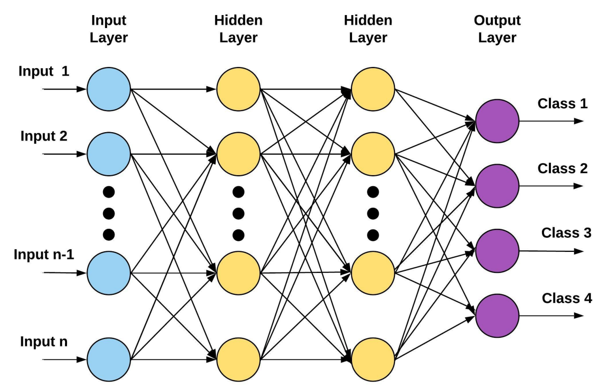

Fig. 10: Multi Layer Perceptron

We have used a feed-forward neural network having an input layer, two hidden layers and an output layer as shown in FIG. 10. We have treated this as a classification problem instead of regression problem because estimating rainfall is an extremely difficult as it depends on a number of factors. Determining the exact rainfall in (mm) requires a lot of corrections which are usually developed and made by experienced scientists.

Architecture details:

- The number of input units are equal to the number of features.
- The number of neurons in each hidden layer is equal to 16*(number of input units).
- The output layer has 4 neurons (one for each class).
- Optimizer - Adam Optimizer
- Loss Function - Categorical cross-entropy

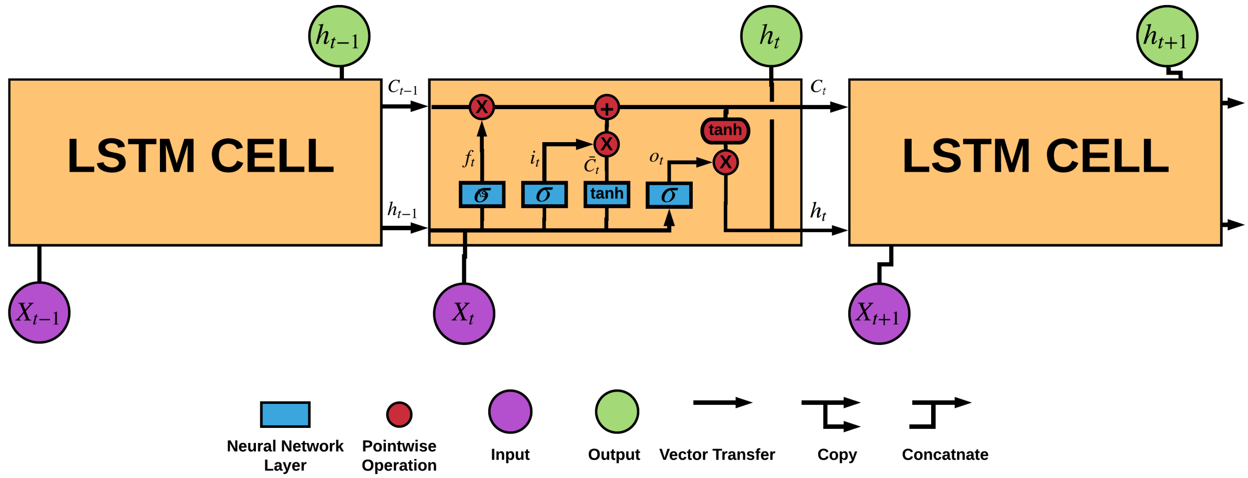

Fig. 11: Long Short Term Memory Module

- Activation Function - relu for hidden layers and softmax for output layer

class ID	Rainfall(in mm)
class 0	0
class 1	<2
class 2	<5
class 3	>=5

TABLE II: Class ID and its corresponding rainfall (in mm) range

B. Synthetic Minority Oversampling Technique (SMOTE)

SMOTE is a synthetic data generation technique which helps in balancing classes. It generates new samples by choosing k neighbours for each existing minority class sample and then estimates some mean value among those k neighbours. This is done continuously till the majority and minority classes have same number of samples.

This is a useful technique for estimating rainfall hourly because of the known fact that it doesn't usually rain every hour and most of the hours there is zero rainfall. Nearly 90% of the data contains zero rainfall (majority class). Due to this imbalance, the classifier always learns to predict the majority class. To avoid this and do the rightful prediction of rainfall we use SMOTE.

C. Long short-term memory (LSTM)

An LSTM network contains LSTM cells also commonly known as the memory units. Each memory cell contains input, output and a forget gate. The LSTM module is shown in FIG. 11 Rainfall varies both spatially and temporally and is almost unpredictable. As we have rainfall values every hour, this is time series data. LSTM works well with sequence data because it can remember what it has seen before (previous time steps) and use that to make its predictions. Sometimes it is useful to know what happens in the next time step and also use that in present time step's prediction. So, we use a Bidirectional LSTM network which captures this behaviour. Time Distributed dense layer helps in creating one to one mapping between input and output because for each time step it connects a dense layer.

Architecture details:

- The number of input units are equal to the number of features.
- First layer is Bidirectional LSTM layer with 24 LSTM cells for dataset 1 and 10 LSTM cells for dataset 2.
- The next layer is a time distributed dense layer.
- In output layer each time step has 4 neurons (one for each class).
- Optimizer - Adam Optimizer
- Loss Function - Categorical cross-entropy
- Activation Function - softmax for output layer

VI. Results

A. Metrics [21]

$$Accuracy = \frac{Number\ of\ correctly\ predicted\ samples}{Total\ number\ of\ sample\ predictions}$$

$$Precision = \frac{Number\ of\ true\ positives}{Number\ of\ true\ positives + Number\ of\ false\ positives}$$

$$Recall = \frac{Number\ of\ true\ positives}{Number\ of\ true\ positives + Number\ of\ false\ negatives}$$

$$Fscore = \frac{2 \times Precision \times Recall}{Precision + Recall}$$

	One Station			Dataset 1			Dataset 2		
	MLP wihout SMOTE	MLP with SMOTE	LSTM	MLP without SMOTE	MLP with SMOTE	LSTM	MLP without SMOTE	MLP with SMOTE	LSTM
categorical_accuracy	0.93	0.91	0.95	0.80	0.5	0.85	0.62	0.46	0.84
val_categorical_accuracy	0.92	0.90	0.95	0.80	0.54	0.83	0.63	0.47	0.84
Precision overall	0.93	0.87	-	0.80	0.66	-	0.64	0.65	-
Recall overall	0.93	0.83	-	0.80	0.37	-	0.60	0.23	-
F-score overall	0.93	0.85	-	0.80	0.47	-	0.62	0.34	-
Precision_class1	0.00	0.79	-	0.50	0.62	-	0.66	0.68	-
Recall_class1	0.00	0.45	-	0.00	0.27	-	0.12	0.16	-
F-score_class1	0.00	0.57	-	0.00	0.38	-	0.20	0.26	-
Precision_class2	0.00	0.79	-	0.00	0.66	-	0.65	0.69	-
Recall_class2	0.00	0.67	-	0.00	0.37	-	0.01	0.16	-
F-score_class2	0.00	0.72	-	0.00	0.47	-	0.02	0.26	-
Precision_class3	0.00	0.83	-	1.00	0.77	-	0.68	0.78	-
Recall_class3	0.00	0.71	-	0.00	0.54	-	0.01	0.25	-
F-score_class3	0.00	0.77	-	0.00	0.63	-	0.01	0.38	-

TABLE III: Results

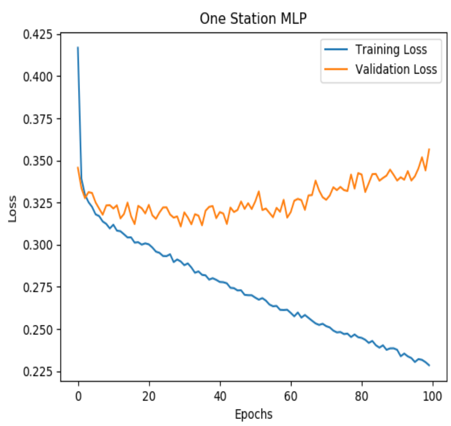

Fig. 12: MLP model for SAC BHOPAL without SMOTE

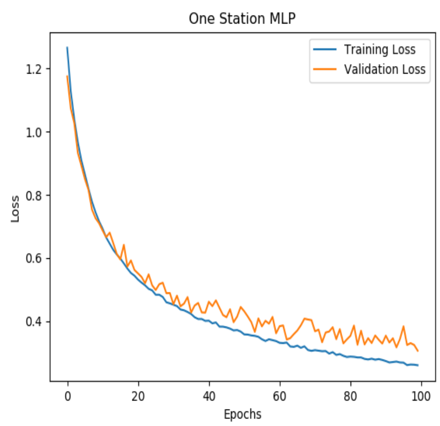

Fig. 13: MLP model for SAC BOPAL with SMOTE

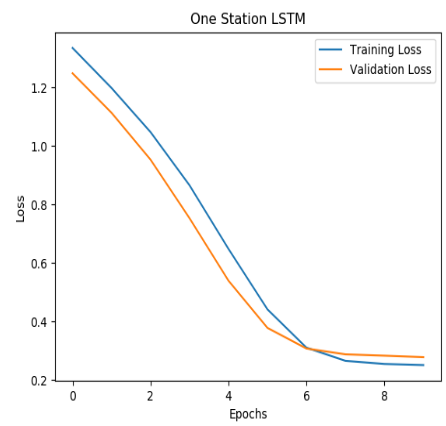

Fig. 14: LSTM model for SAC BOPAL

B. Analysis

In FIG. 12 the training loss decreases but the validation almost stays the same (with minimal changes). This happens due to high imbalance in classes. The model always predicts 0 giving it a very high accuracy of 92.17% but in reality the model performs poorly. This is justified by the classes 1, 2 and 3 having f-score of 0 as seen in TABLE III. On generating synthetic data we observe that over multiple epochs the model learns with both the training and validation loss decreasing as seen in FIG. 13. In this case we get an accuracy of 89.47% and the model is robust as it predicts all the classes (not just favouring one class). In this case the classes have a good f-score (almost same for all classes). Looking at the loss curves in FIG. 14 it seems that the classifier is learning a lot but it just predicts 0 in all cases. High accuracy makes no sense in this case. We need balanced classes for better results.

When the model was trained for entire state of Gujarat, our observations were similar to the case of one station. Without SMOTE, training loss decreases whereas the validation loss remains the same as seen in FIG. 15. The precision values were comparatively better because of increasing data. In FIG. 16, we see that the loss curves are decreasing and the model is more generalized as compared to the one with one station. The following metric (precision, recall and f-score) values suggest that the model is not biased to any class. On increasing the data, the LSTM performs better and doesn't always predict 0. In FIG. 17, the learning curves imply the models fast learning rate. As it was trained for entire Gujarat state we get a better model.

In dataset 2 most of the samples belong to class 0 and class 1 as seen in FIG. 21. Thus MLP model without smote gives comparatively good f-score of 0.20 for class 1 and very less (0.01) for class 2 and class 3. In FIG. 18 though the loss curves are decreasing, this model isn't good. We get decreasing curves in FIG. 19 with slight variations but this model generalizes well with decent accuracy. LSTM model for dataset 2 works the best among all with a good accuracy of 84% even though there were only 60% zeros in the training label set. This is good indication of the model not just predicting zeros but the actual rainfall. Refer to FIG. 20 for the learning curves.

VII. Future Work

- We have currently used only 2014 year data for preparing dataset 2 and training the models due to limitation of computing resources. Given advanced computing settings we expect to get better results.

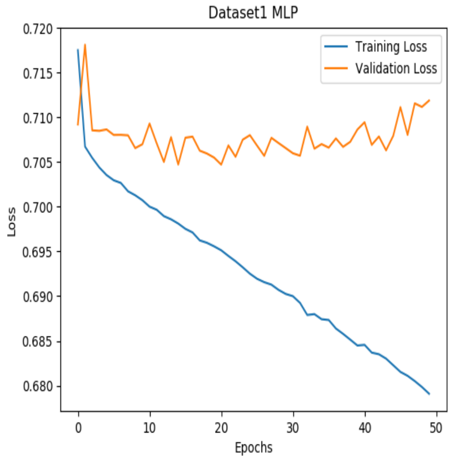

Fig. 15: MLP model for Gujarat without SMOTE using Dataset 1

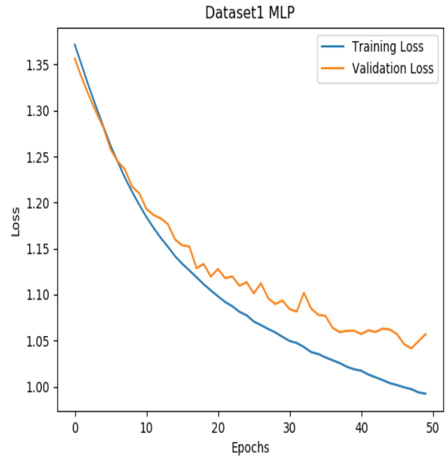

Fig. 16: MLP model for Gujarat with SMOTE using Dataset 1

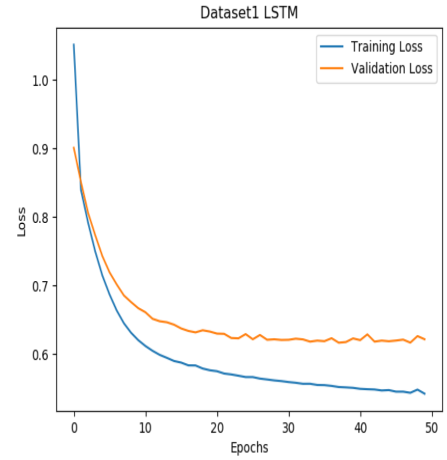

Fig. 17: LSTM model for Gujarat using Dataset 1

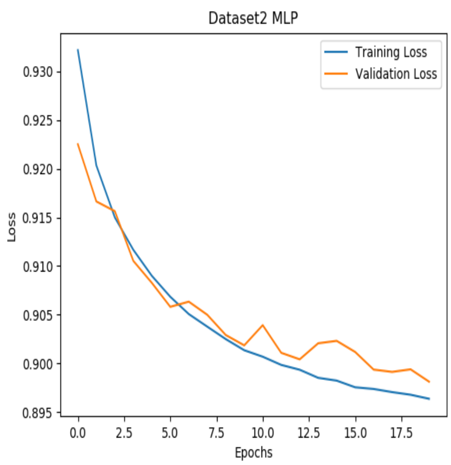

Fig. 18: MLP model for Gujarat without SMOTE using Dataset 2

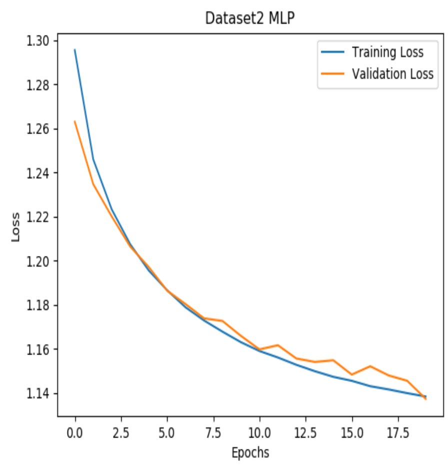

Fig. 19: MLP model for Gujarat with SMOTE using Dataset 2

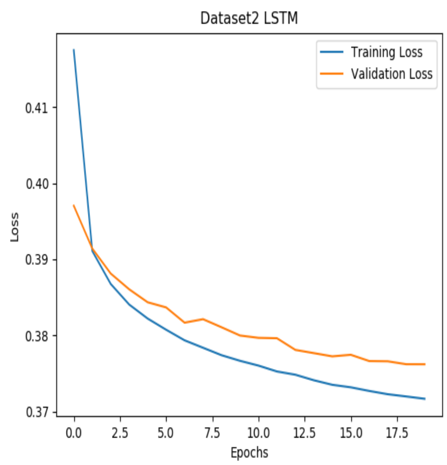

Fig. 20: LSTM model for Gujarat using Dataset 2

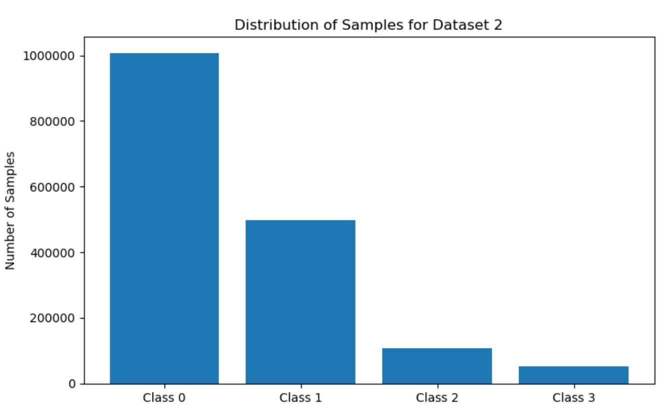

Fig. 21: Sample Distribution for Dataset 2

- In this work we have only focused on Gujarat, we would like to develop this model for entire India.
- Though we have adopted the best possible models for predicting rainfall (after detailed literature review), we could explore some more models.
- We believe that the IR satellite data taken from MOSDAC to be not entirely accurate so we would try to incorporate data provided by NASA as well.
- Addition of PMW (Passive Microwave) can be useful as it is directly related to the water content in the clouds but the payload of PMW is carried only on the low earth orbit. Due to it's swath scanning procedure it takes several hours to completely scan the earth, thus we have to add necessary corrections to generate intermediate data.
- We can incorporate more spectral bands and their corresponding texture features in the input.

VIII. Conclusion

This paper discusses the importance of different parameters that determine the amount of rainfall occurring in a particular region. Due to limitation in computing resources, we could only train the models for the year of 2014 (while using dataset 2). Using multispectral channels (in dataset 2) over TIR1 channel (in dataset 1) didn't improve the results significantly as stated in some papers. So direct comparison between data-sets could not made with substantial evidence. MLP gave good results when SMOTE technique was used for a particular station data (small dataset). The metric values (f-score, precision, recall) for each class can as seen in TABLE III. This shows that the model predictions have no bias and is generalized. LSTM performed better than MLP in general as it made predictions with an accuracy of 84% when dataset 2 was used (relatively balanced than dataset 1).

Acknowledgment

We thank Mr. Arjun Bhasin, Head of Research and Development, Amnex InfoTechnologies Private Limited, for providing us with the opportunity to work on this project. We appreciate his constant support and guidance. We also thank Shivani Shah, Senior Scientist at ISRO, SAC for her technical assistance on acquisition and understanding of INSAT-3D data.

References

[1] Phillip A Arkin and Bernard N Meisner. "The relationship between large-scale convective rainfall and cold cloud over the western hemisphere during 1982-84". In: *Monthly Weather Review* 115.1 (1987), pp. 51–74.

[2] Christian Kummerow, Y Hong, WS Olson, S Yang, RF Adler, J McCollum, R Ferraro, G Petty, Dong-Bin Shin, and TT Wilheit. "The evolution of the Goddard Profiling Algorithm (GPROF) for rainfall estimation from passive microwave sensors". In: *Journal of Applied Meteorology* 40.11 (2001), pp. 1801–1820.

[3] Robert J Joyce, John E Janowiak, Phillip A Arkin, and Pingping Xie. "CMORPH: A method that produces global precipitation estimates from passive microwave and infrared data at high spatial and temporal resolution". In: *Journal of Hydrometeorology* 5.3 (2004), pp. 487–503.

[4] Martin C Todd, Chris Kidd, Dominic Kniveton, and Tim J Bellerby. "A combined satellite infrared and passive microwave technique for estimation of small-scale rainfall". In: *Journal of Atmospheric and Oceanic Technology* 18.5 (2001), pp. 742–755.

[5] Christian Kummerow and Louis Giglio. "A method for combining passive microwave and infrared rainfall observations". In: *Journal of Atmospheric and Oceanic Technology* 12.1 (1995), pp. 33–45.

[6] F Joseph Turk, Elizabeth E Ebert, Hyun-Jong Oh, Byung-Ju Sohn, Vincenzo Levizzani, Eric A Smith, and Ralph Ferraro. "Validation of an operational global precipitation analysis at short time scales". In: *12th Conference on Satellite Meteorology and Oceanography and 3rd Conference on Artificial Intelligence Applications to Environmental Science, Seattle, Washington*. 2003.

[7] Patrick WS King, William D Hogg, and Philip A Arkin. "The role of visible data in improving satellite rain-rate estimates". In: *Journal of Applied Meteorology* 34.7 (1995), pp. 1608–1621.

[8] Tim Bellerby, Martin Todd, Dom Kniveton, and Chris Kidd. "Rainfall estimation from a combination of TRMM precipitation radar and GOES multispectral satellite imagery through the use of an artificial neural network". In: *Journal of applied Meteorology* 39.12 (2000), pp. 2115–2128.

[9] Toshiyuki Kurino. "A satellite infrared technique for estimating deep-shallow precipitation". In: *Advances in Space Research* 19.3 (1997), pp. 511–514.

[10] Mamoudou B Ba and Arnold Gruber. "GOES multi-spectral rainfall algorithm (GMSRA)". In: *Journal of Applied Meteorology* 40.8 (2001), pp. 1500–1514.

[11] Robert J Kuligowski. "A self-calibrating real-time GOES rainfall algorithm for short-term rainfall estimates". In: *Journal of Hydrometeorology* 3.2 (2002), pp. 112–130.

[12] Ali Behrangi, Kuo-lin Hsu, Bisher Imam, Soroosh Sorooshian, George J Huffman, and Robert J Kuligowski. "PERSIANN-MSA: A precipitation estimation method from satellite-based multispectral analysis". In: *Journal of Hydrometeorology* 10.6 (2009), pp. 1414–1429.

[13] *MOSDAC INSAT-3D, Indian Space Research Organization(ISRO)*. https://www.mosdac.gov.in.

[14] Daniel Rosenfeld and Garik Gutman. "Retrieving microphysical properties near the tops of potential rain clouds by multispectral analysis of AVHRR data". In: *Atmospheric research* 34.1-4 (1994), pp. 259–283.

[15] CM Kishtawal. "Meteorological satellites". In: *Satellite Remote Sensing and GIS Applications in Agricultural Meteorology* 73 (2005).

[16] Albert Arking and Jeffrey D Childs. "Retrieval of cloud cover parameters from multispectral satellite images". In: *Journal of Climate and Applied Meteorology* 24.4 (1985), pp. 322–333.

[17] Ali Behrangi, Kuo-lin Hsu, Bisher Imam, Soroosh Sorooshian, and Robert J Kuligowski. "Evaluating the utility of multispectral information in delineating the areal extent of precipitation". In: *Journal of Hydrometeorology* 10.3 (2009), pp. 684–700.

[18] M Cheng, R Brown, and CC Collier. "Delineation of precipitation areas using Meteosat infrared and visible data in the region of the United Kingdom". In: *Journal of Applied Meteorology* 32.5 (1993), pp. 884–898.

[19] Andy Jarvis, Hannes Isaak Reuter, Andrew Nelson, Edward Guevara, et al. "Hole-filled SRTM for the globe

Version 4”. In: *available from the CGIAR-CSI SRTM 90m Database (http://srtm. csi. cgiar. org)* 15 (2008).

[20] *Normalized Difference Vegetation Index (NDVI), SDAPSA, National Remote Sensing Center, Bhuvan Noeda*. http://bhuvan.nrsc.gov.in.

[21] Jesse Davis and Mark Goadrich. “The Relationship Between Precision-Recall and ROC Curves”. In: *Proceedings of the 23rd International Conference on Machine Learning*. ICML ’06. Pittsburgh, Pennsylvania, USA: ACM, 2006, pp. 233–240. ISBN: 1-59593-383-2. DOI: 10.1145/1143844.1143874. URL: http://doi.acm.org/10.1145/1143844.1143874.

List of Abbreviations

AWIFS	Advanced Wide Field Sensor
AWS	Automatic Weather Station
DEM	Digital Elevation Models
LSTM	Long-Short term Memory
MIR	Mid wave Infrared (0.39μm - 0.7μm)
MLP	Multi Layer Perceptron
NDVI	Normalized Difference Vegetation Index
PMW	Passive Microwave
SRTM	Shuttle Radar Topography Mission
SWIR	Short Wave Infrared (1.55μm - 1.70μm)
SZA	Solar Zenith Angle
TIR1	Thermal Infrared (10.2μm - 11.2μm)
TIR2	Thermal Infrared (11.5μm - 12.5μm)
TRMM	Tropical Rainfall Measuring Mission
VIS	Visual Wavelength
WV	Water Vapor Wavelength (6.5μm - 7.0μm)