

Predicting Composition of Photo Voltaic Cells Using Neural Networks

Introduction: A better understanding of IV curve data collected from photo voltaic cells may lead to the construction of better solar cells. With this in mind IV curve data from different types of solar cells was acquired from Photovoltaics and Electrochemical Systems Branch, NASA Glenn Research Center. Neural networks were created to predict the chemical composition of different classes of solar cells with varying degrees of success.

Language and API: The neural network was built in Python using Keras with TensorFlow as a backend, and the data to train the network was acquired from researchers from NASA Goddard Research Center. Data preprocessing steps and K-means clustering were conducted using Sci-Kit Learn, and electrical properties of the curve were calculated using R.

Solar Cell Data: The IV curves from over 7000 solar cells representing 20 different types of solar cell materials were used in this study. The short circuit current (SCC), reverse saturation current (RCS), shunt resistance (RSH), open circuit voltage (VOC), series resistance RS, and the diode ideality (DIF) were calculated from the IV curves using a proprietary R code developed at NASA Goddard. These values were used as the features to predict the materials used in the different solar cells. The solar data was first separated by material type, which were non-alloy compound and elemental, single gap, and triple gap solar cells. Non-alloy compound and elemental solar cells were those composed of elements or non-alloy compounds. Single gap solar cells are those which possess only one energy gap, and triple gap cells are those that possess three energy gaps. There were 5 non-alloy compound and elemental solar cells, 15 single gap solar cells and 3 triple gap solar cells.

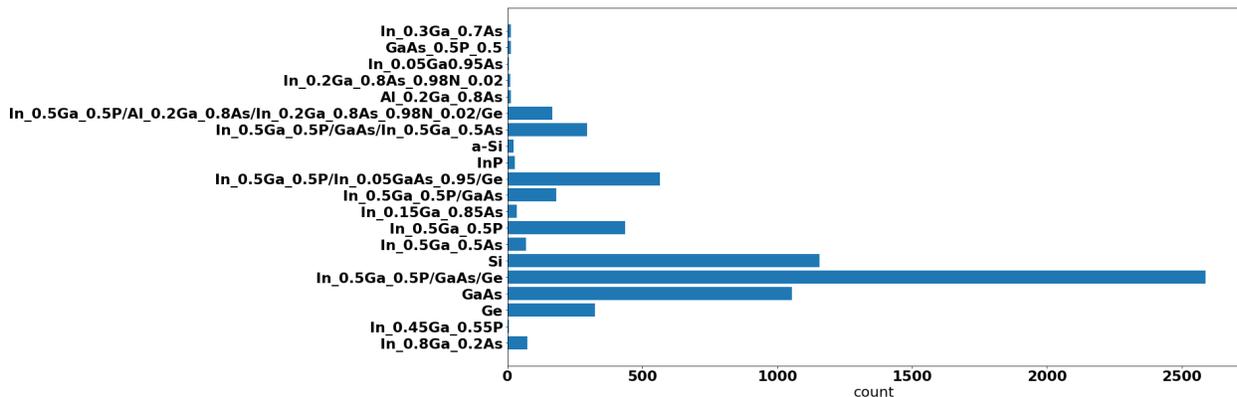


Figure 1: The y-axis is listing of solar cell materials represented in the received data set. The x-axis is the count of a particular material.

K-mean clustering of the PV cell IV curve data suggests that there are four types of PV cell IV curves, which are visualized in the t-SNE plot (Figure 2a). Interestingly, each curve cluster has a unique chemical signature (Figure 2b), which suggests that chemistry plays an important role in

determining solar cell performance. However, IV curves do not cluster solely on the chemical properties that were used to divide the compounds to train neural network models.

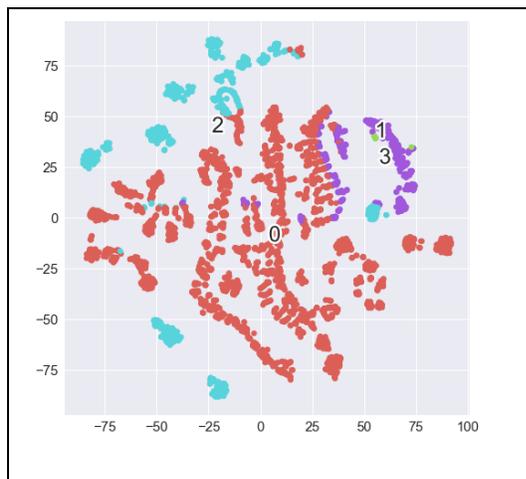


Figure 2a: t-SNE visualization of the 4 k-means clusters. Each point represents an IV curve and each color represents a distinct cluster (0 – red, 1 – green, 2 – cyan, 3 – purple). The coordinates are for reference only and are not to be interpreted as a distance relationship among the data.

TEXT	Coordinate	
0	4.02827, -11.7385	
1	55.2538, 39.4362	
2	-23.388, 41.9674	
3	57.2296, 26.5328	

Cluster	Number of Curves	
0	4693	
1	8	
2	1188	
3	652	


```
def nn_model(out, l1, l2, l3):
    model = Sequential()
    model.add(Dense(l1, input_shape = (6,), kernel_initializer = glorot_uniform(seed=None), activation = 'relu'))
    model.add(Dropout(0.1))
    model.add(Dense(l2, kernel_initializer = glorot_uniform(seed=None), activation = 'relu'))
    model.add(Dropout(0.1))
    model.add(Dense(l3, kernel_initializer = glorot_uniform(seed=None), activation = 'relu'))
    model.add(Dropout(0.1))
    model.add(Dense(out, activation = 'softmax'))
    return model, es
```

Code
Keras Code for the
Neural Network

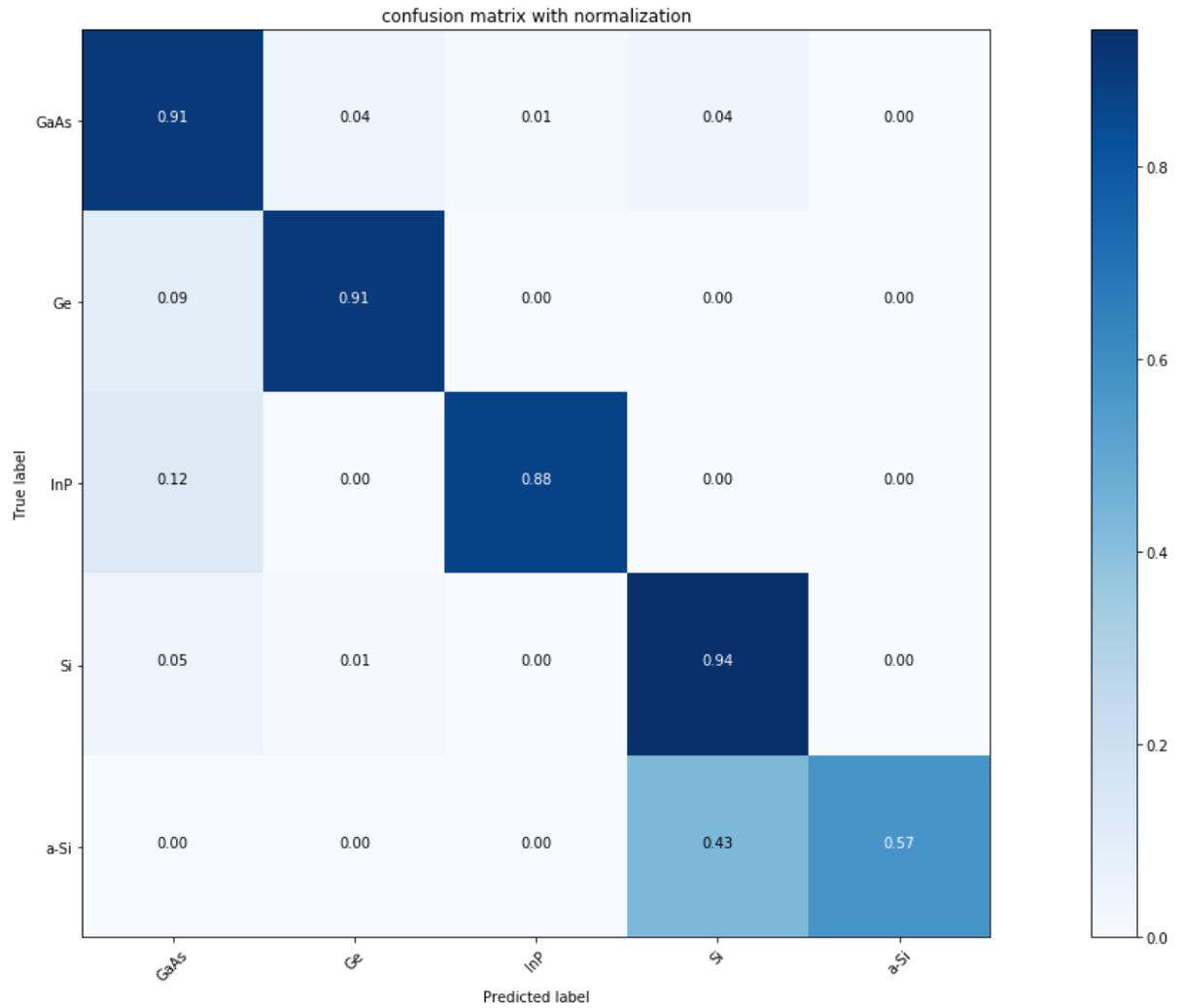
Layer (type)	Output Shape	Param #
dense_1 (Dense)	(None, 50)	350
dropout_1 (Dropout)	(None, 50)	0
dense_2 (Dense)	(None, 30)	1530
dropout_2 (Dropout)	(None, 30)	0
dense_3 (Dense)	(None, 20)	620
dropout_3 (Dropout)	(None, 20)	0
dense_4 (Dense)	(None, 5)	105
Total params: 2,605		
Trainable params: 2,605		
Non-trainable params: 0		

Dense
Linear Model of
the Data

Predicted Composition(s)

Figure 3: Keras code for the fully connected neural network.

Results: Of the three models the model distinguishing among the non-alloy compound and elemental solar cells performed the best. With a 91% accuracy it could predict which of the 5 solar cell materials generated a given IV curve given the calculated electrical properties (Figure 3a). The model had the most difficulty distinguishing between a-Si and Si; when presented with calculated electrical data from a given a-Si IV curve it predicted it was Si 43% of the time. However, when presented with calculated electrical data from a given Si IV Curve it predicted it was a Si 94% of the time. The single gap model fared well, correctly predicting the composition 81% of the time (Figure 3b). There were a few materials where the prediction was accurate 100% of the time. There were also a few materials, one of which was a-Si, where it had a prediction accuracy of 0%. The triple gap model fared the worst (Figure 3c). This is likely due to the disproportionate representation of In_{0.5}Ga_{0.5}P/GaAs/Ge in the data set. Consequently, the model defaulted to predicting In_{0.5}Ga_{0.5}P/GaAs/Ge, failing to correctly predict the other compounds the majority of the time.



Figurer 4a: The confusion matrix for the non-alloy compound and elemental solar cells targeted in the model. The number of IV curves from a given chemistry is in parentheses: GaAs (1055), Ge (323), InP (27), Si (1156), a-Si (22)

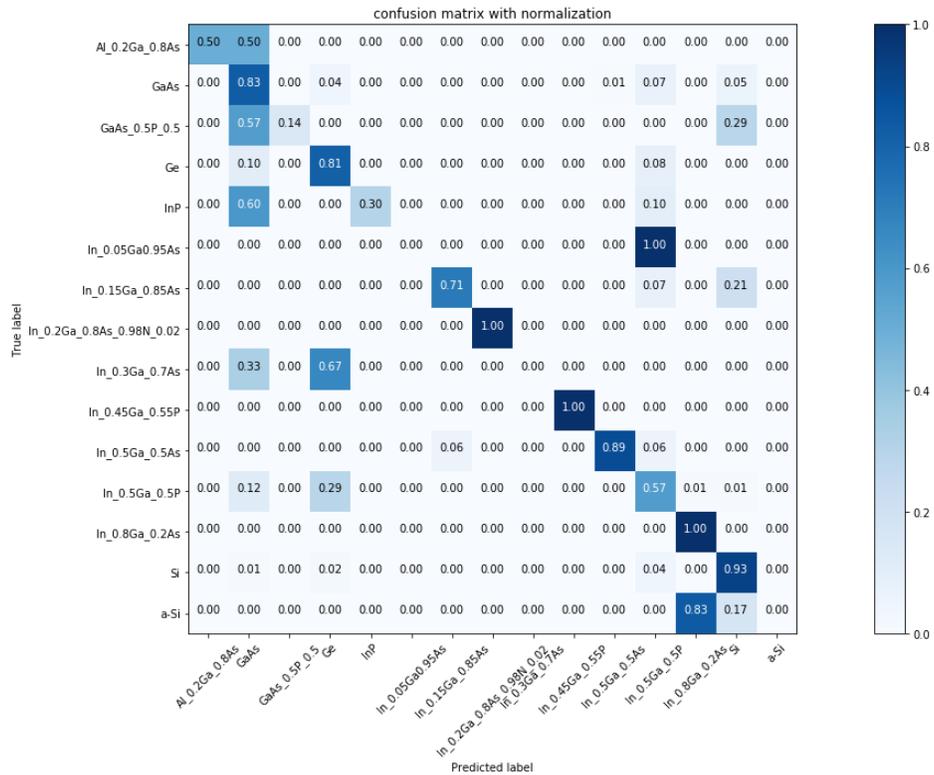


Figure 4b: The confusion matrix for the single gap solar cells targeted in the model. The number of IV curves from a given chemistry is in parentheses: Al_{0.2}Ga_{0.8}As (11), GaAs (1055), GaAs_{0.5}P_{0.5} (12), Ge (323), InP (27), In_{0.05}Ga_{0.95}As (5), In_{0.15}Ga_{0.85}As (34), In_{0.2}Ga_{0.8}As_{0.98}N_{0.02} (9), In_{0.3}Ga_{0.7}As (12), In_{0.45}Ga_{0.55}P (4), In_{0.5}Ga_{0.5}As (67), In_{0.5}Ga_{0.5}P (435), In_{0.8}Ga_{0.2}As (73), Si (1156), a-Si (22)

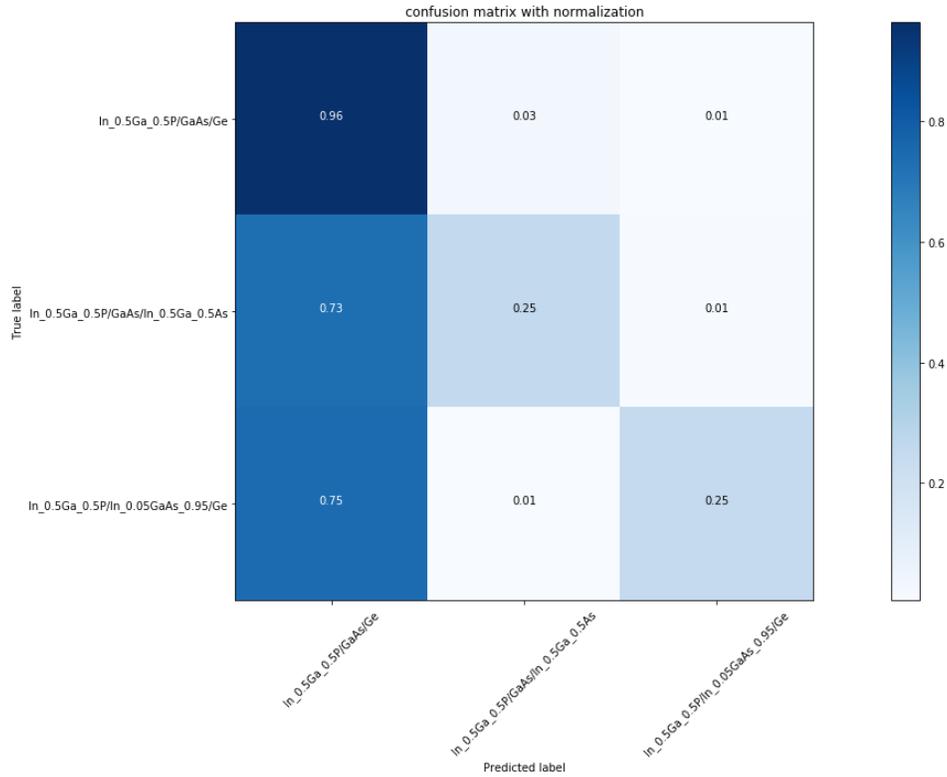


Figure 4c: Confusion matrix for the triple gap solar cell compounds_targeted in the model. The number of IV curves from a given chemistry is in parentheses: In_0.5Ga_0.5P/GaAs/Ge (2588), In_0.5Ga_0.5P/GaAs/In_0.5Ga_0.5As (294), In_0.5Ga_0.5P/In_0.05GaAs_0.95/Ge (565)

Conclusion: Trained neural network models were used to predict the materials composition of solar cells from electrical parameters generated from IV curves. The model worked best predicting non-alloy compound and elemental solar cells. The model prediction was at 91% accuracy for the best-case scenario. More data in equal amounts from the different material types is needed to improve model performance for other chemical compositions.