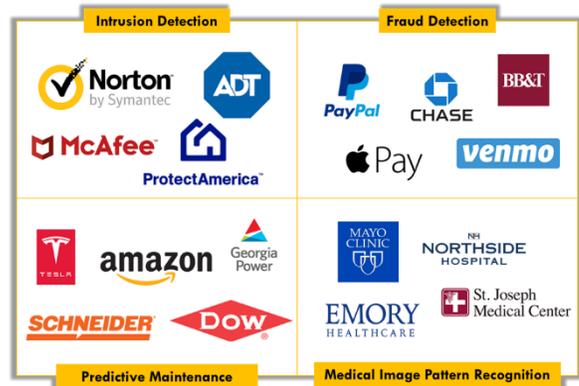


With the extensive quantity of data that businesses have collected to date with regards to customer related information, there are a lot of exciting opportunities from a data science / analytics standpoint with regards to anomaly detection. This is the field of study that attempts to identify rare observations in data that differ significantly from the majority of the data.

In the past, anomaly detection was considered to be a “preprocessing” step that was used to exclude data which could potentially exacerbate model effectiveness. However, as data has grown in size and unpredictability, more attention has been devoted to predicting anomalies as a proactive measure, as opposed to a reactive approach (to remove outliers for modeling purposes). In general, industries have become more aware of drawing the line between ‘business as usual’ variability and anomalous behavior. Aside from this, we can challenge conventional thinking of what is considered as an outlier. Traditionally a data point is considered as an outlier if it is associated to an abnormally high / low response value; however, in reality, that may not be the case. Now we can truly identify if a response value is abnormal or not for a given feature set.

Generally, the major domains of anomaly detection are within software applications (intrusion detection), finance (fraud detection), malfunctioning equipment (predictive maintenance) and medicine (medical image pattern recognition). In any case, the idea is to develop models that will:

1. Identify outliers in the existing data
2. Elucidate an understanding for the drivers that determine these outliers
3. Use the models to predict when an anomaly will occur in the future



## **Autoencoder Neural Networks**

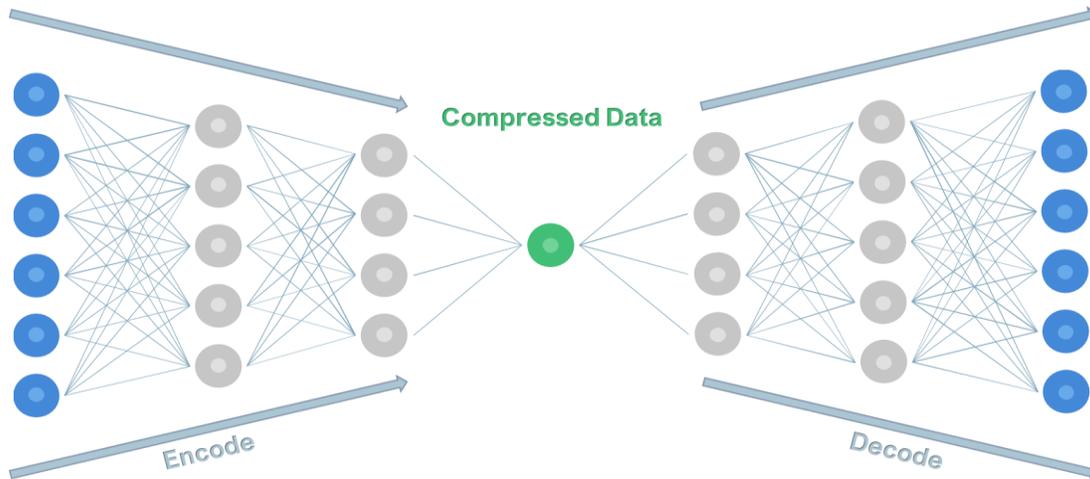
In the current state, there are many machine learning algorithms that attempt to detect outliers i.e. clustering classifiers, support vector machines etc. However recently, there have been more advanced deep learning / artificial intelligence methods that achieve this task more effectively.

An autoencoder neural network is a unique type of neural network that determines the core features from the input and tries to recreate the input itself. This is different from a conventional neural network that outputs a class from a given input. The idea behind this is to be able to encode the input data into key features that are most influential in determining if a point is considered normal or not. An important point to note during implementation is that the model is only trained on normal data - in other words, the subset of the training split data that is within "normal" class is used to train the model. This in turn identifies any new data as either normal or not (i.e. an anomaly).

The idea behind autoencoders is that it has symmetric encoding and decoding layers, which have a lower dimension than the input and output dimension. As the autoencoder tries to map its input into its output, the input and output layers both have the same dimensions. As the data is pushed from 1 layer to the next within the encoding phase, it is important to note that the number of dimensions in each layer reduces.



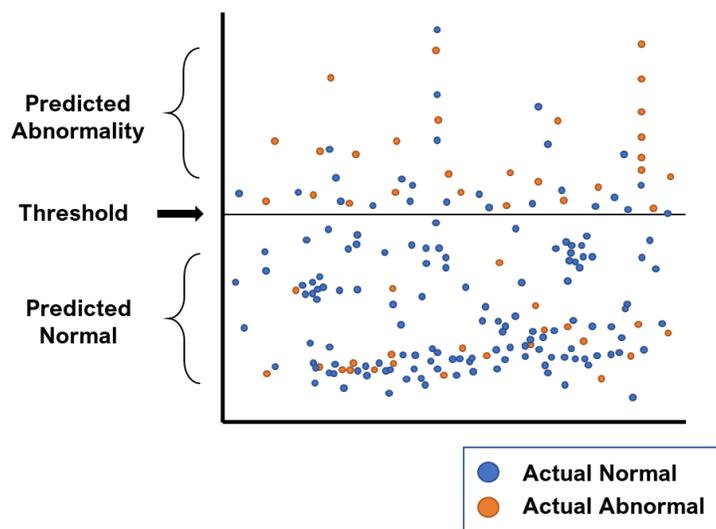
Conversely, the opposite should occur as data is pushed from 1 layer to the next within the decoding phase. This has been represented by the diagram below:



Within autoencoders, the concept of reconstruction error (RE) is introduced - the Euclidean distance between the input and its mapped output. The RE value is a comparative measure between the actual input data as well as its associated predicted data from the model. Any large RE values that deviate from most of the other RE values are considered to be outliers. For a given threshold, any points with RE values greater than the threshold are considered to be outliers.

### **Feature Contribution to Errors**

For every data point, each RE per point can be broken down into RE per feature. Since the RE is the Euclidean distance between the initial point and the predicted point, the error per dimension can be calculated by undertaking coalition analysis. Now, the error for that row is in terms of every feature. Doing this process for all the outlier points and summing the RE per dimension for these points allows us to account for the features that are most influential towards the determination of outliers. Taking the total RE values against every feature allow us to see which features contribute to the largest RE.



*Kaizen Thought Leadership:*  
**A NEW APPROACH TO ANOMALY DETECTION**

---

Finally, the effectiveness of this model can be determined by calculating a confusion matrix. Now, we have established that the initial actual data was divided into 2 classes i.e. normal and abnormal. We can do the same method for the predicted classes based on the RE values. The idea behind this reflects the notions of true positives and negatives, and false positives and negatives i.e. it determines the number of points that the model correctly / incorrectly predicts as normal, and correctly / incorrectly as abnormal.

### ***Applications***

One of the general applications of anomaly detection is within companies' existing user data. For example, within the hospitality industry, this can be used to identify anomalous restaurant transactions against a feature space consisting of weather-related information, public holidays, seasonality variables and reservations-based data. Alongside **KaizenDataLabs™**, these models have the capacity to provide extremely meaningful insights to the internal business as well as the client base.

**About the Authors:** [Adi Banerjee](#) is a Data Scientist at Kaizen Analytix based out of the Atlanta office

