Linear Separability of any Data Projected by a Nonlinear Function

Alok Jadhav, U1265865 Ambuj Arora, U1265867

October 2019

Projecting any dataset from a lower dimension to a higher dimension is the underlying principle of many machine learning algorithms, like SVM, ANNs. SVM uses a kernel function to project lower dimensional-nonlinear data into a much higher dimension so that the data becomes linearly separa-ble. [4] Similarly, in Neural Networks, each layer transforms/ converts lower-dimensional data into higher dimensions. A combination of such layers makes learning possible in complex non-linear datasets.

It is known that any low-dimensional non-linear data is linearly separable in a higher dimension [1]. Thus, as long as we have infinite time and compu-tational power, we are guaranteed to find a dimension in which the training data is linearly separable. The state of the art method is to try and fit dif-ferent transformation functions (φ) on the data to make it linearly separable in any higher dimension. But we do not have infinite resources to achieve that. Also, when we use a transformation function on the data, we can not be sure whether the data is linearly separable in that space or not until we actually try to learn the data. [2]

Our proposed study will try to come up with a mathematical logic such that given a non-linear transformation function and a dataset, we can be sure that the dataset will be linearly separable or not in the higher space without having to try to learn the data. This project will be useful in the field of Liquid state Machines [3] where researchers are trying to harness non-linearity from natural phenomenons (like electromagnetic interference, uncertainty of photons, etc.) to project data from lower dimension to higher dimension. We can also extend this method to measure the effectiveness of ’transformation functions’ in SVM and ANNs.

We plan to start with 2-D space and a data that is non-linear in 2-D. We will then try to make the data linear in 3-D using a non-linear transformation function and try to come up with a mathematical logic that is dependent on the transformation function and the data, which governs the linear separability of the data in 3-D. We will then try to generalize this logic for any higher dimension.

References

[1] Igor Carron et.al Alaa Saade, Francesco Caltagirone. Random projections through multiple optical scattering: Approximating kernels at the speed of light. Arxiv, 2015.

[2] et.al Degang Chena, Qiang Heb. On linear separability of data sets in feature space. Science Direct, 2007.

[3] Jean Benoit H´eroux el al Gouhei Tanaka, Toshiyuki Yamane. Recent advances in physical reservoir computing: A review. 2018.

[4] Kevin Murphy. Machine learning: A probabilistic perspective, 2012.