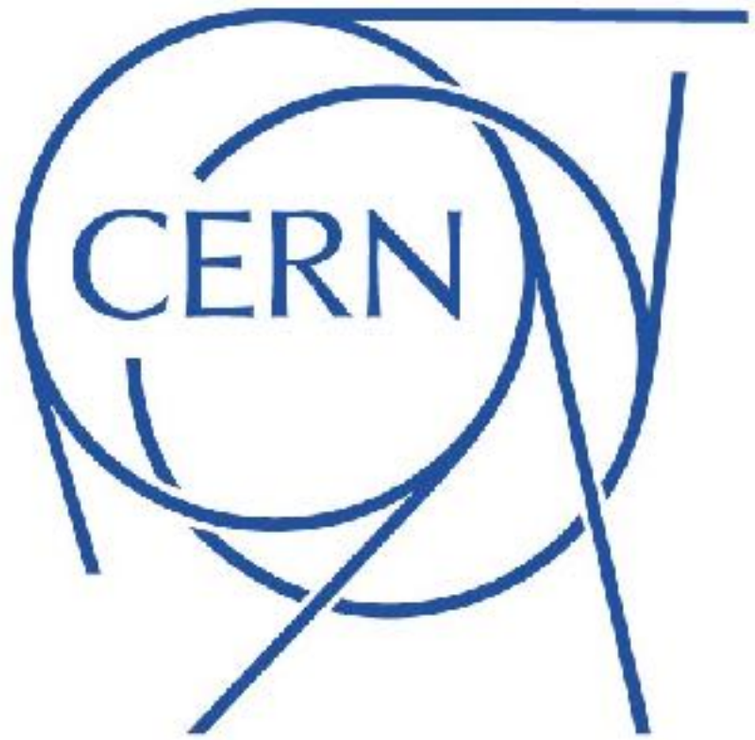


TMVA Project in Machine Learning

Abhinav Moudgil

Mentors

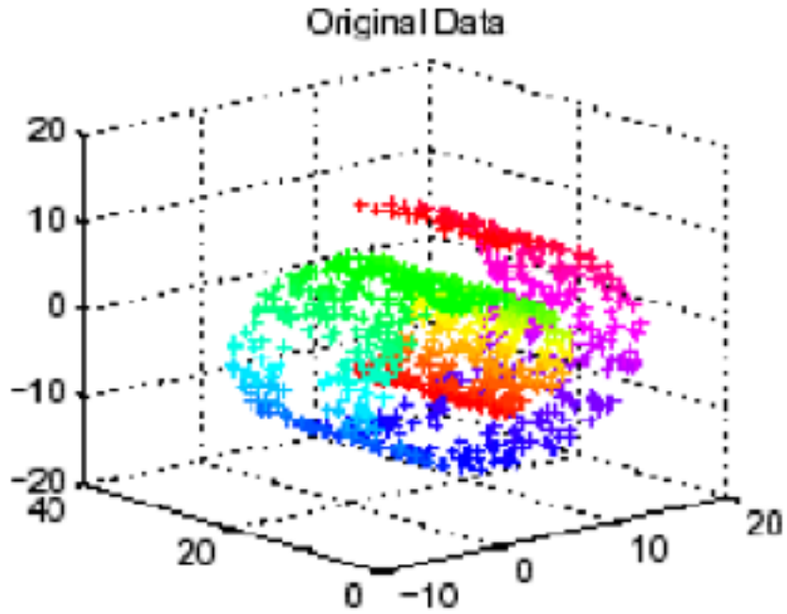
Sergei Gleyzer

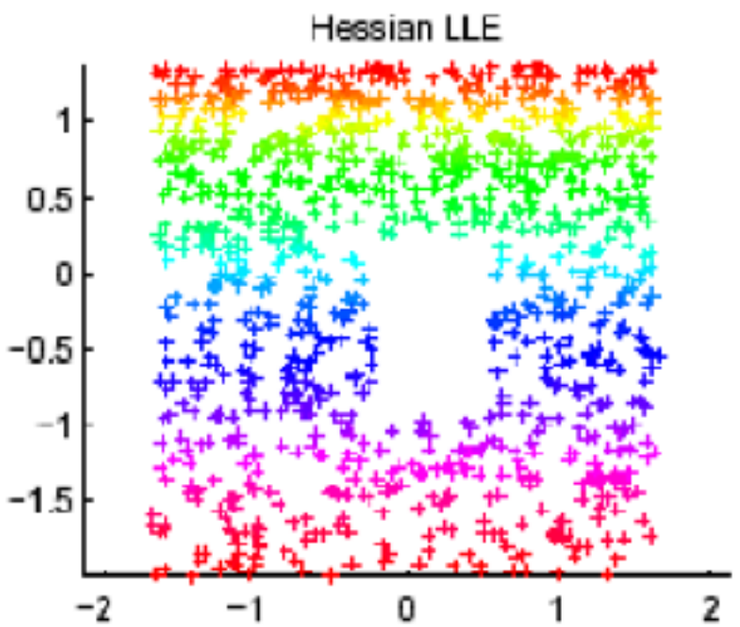
Lorenzo Moneta

Unsupervised Feature Learning

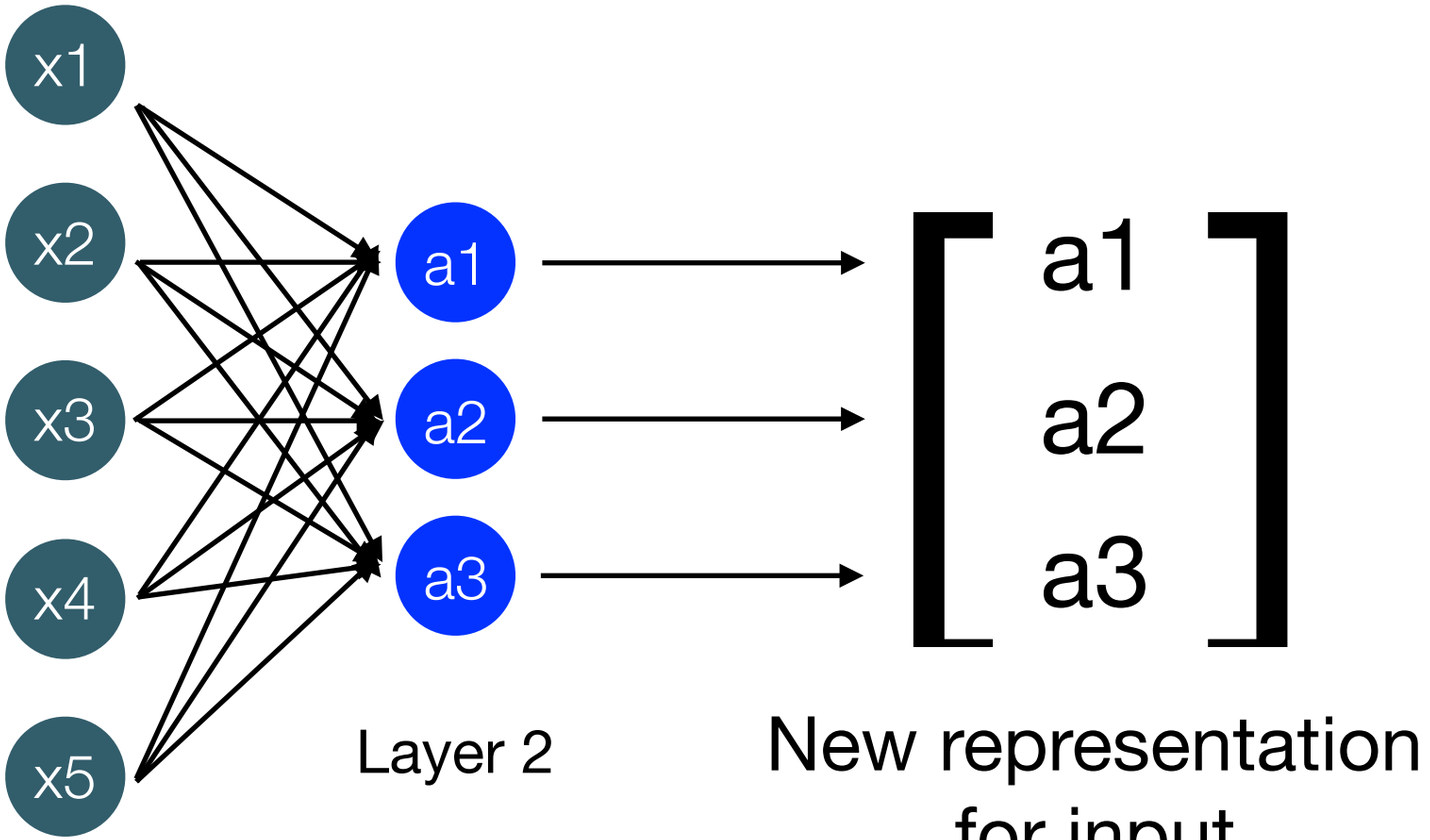
- Transformation of raw data input to a representation that can be effectively exploited in machine learning tasks

- New framework for variable transformations in TMVA with VarTransformHandler class





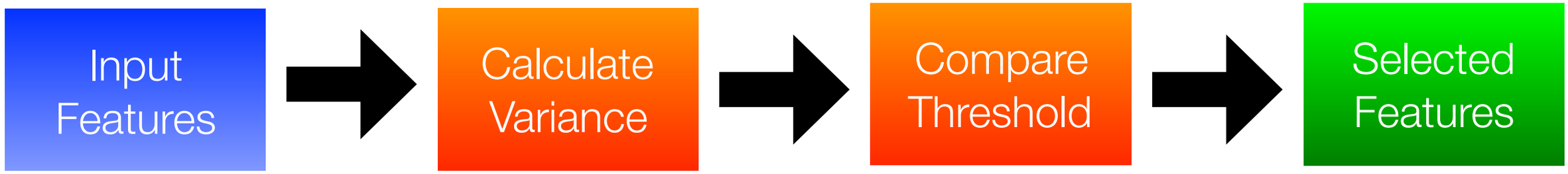
Deep Autoencoders



- Deep neural network is trained to output the input i.e. learn the identity functions.
- Constrain number of units in hidden layer, thus learning compressed representation.

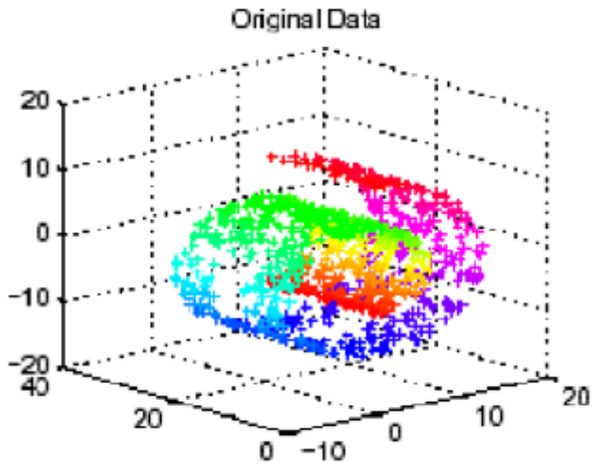
Layer 1

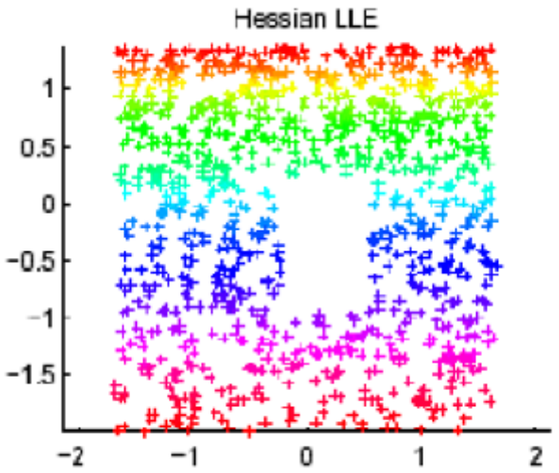
Variance Threshold



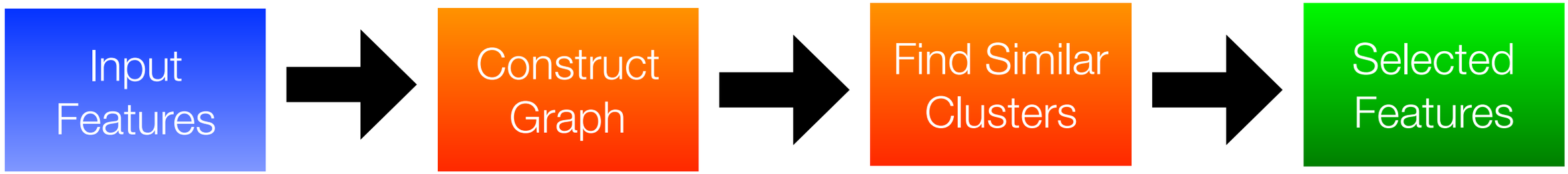
Hessian Local Linear Embedding

- A non linear dimensionality reduction method
- Embeds a set of points from high dimensional space to low dimensional space such that projected point should have the same neighbourhood as the original point





Feature Clustering



VarTransformHandler class

- A new C++ class for all unsupervised feature extraction methods
- VarTransformHandler object takes the user inputs
- Returns new DataLoader with transformed variables

User Interface in TMVA

TMVA::DataLoader* newLoader = loader->VarTransform(option_string)	
Variance Threshold	“VT(threshold_value)”
Deep Autoencoders	“AE(network_layout)”
Feature Clustering	“FC(number_of_dimensions)”
Hessian LLE	“HLLE(number_of_dimensions, number_of_neighbours)”

Status

- Variance Threshold with VarTransformHandler Class
  Merged in TMVA

- Deep Autoencoders
  Prototype developed which works on earlier DNN implementation

- HLLE, Feature Clustering
  Partially implemented due to large memory requirement constraint

Thank You.