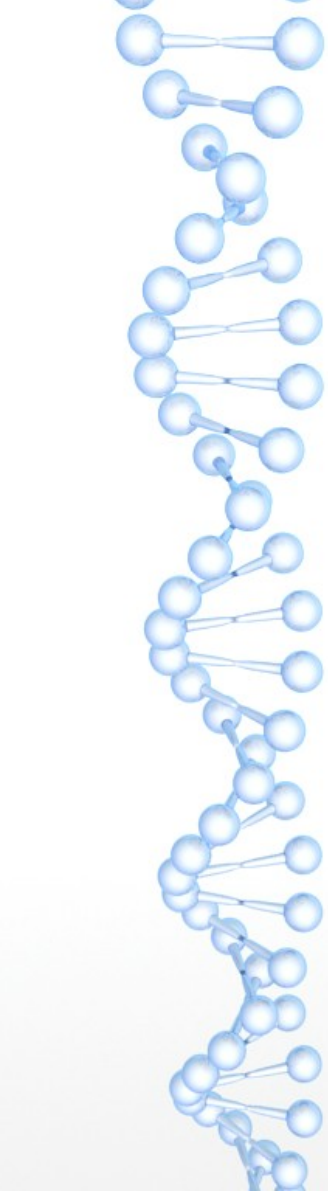
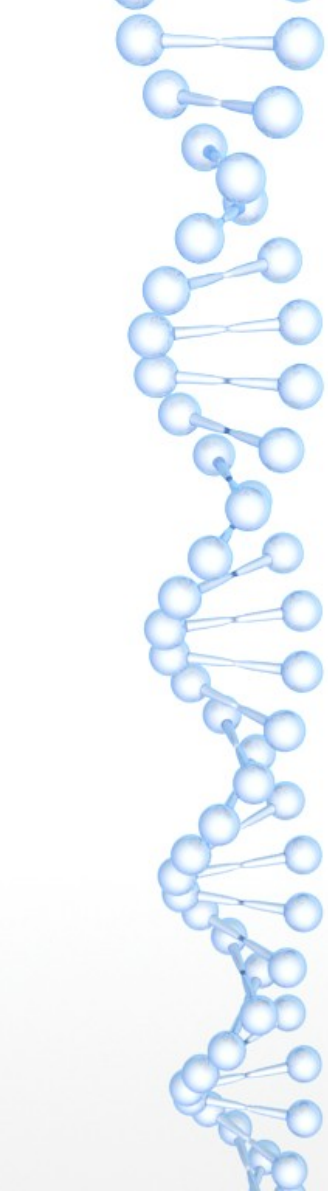


TensorFlow: Biology's Gateway to Deep Learning?

Ladislav et al.

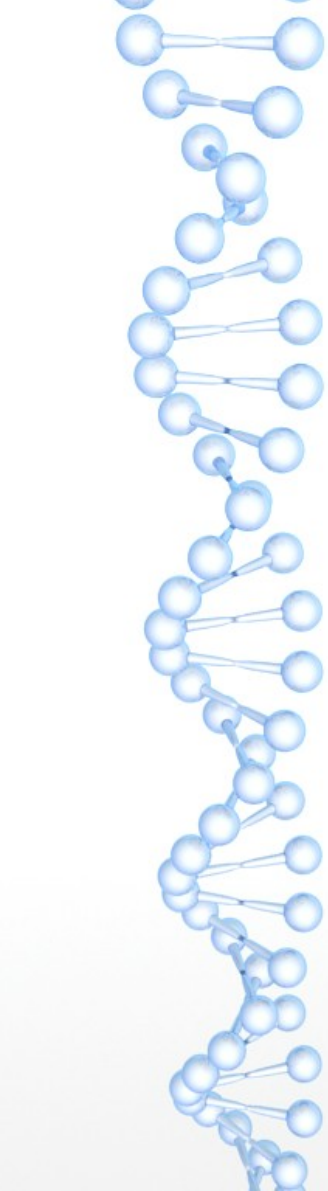
Presented by Wilson Mudaki



Introduction

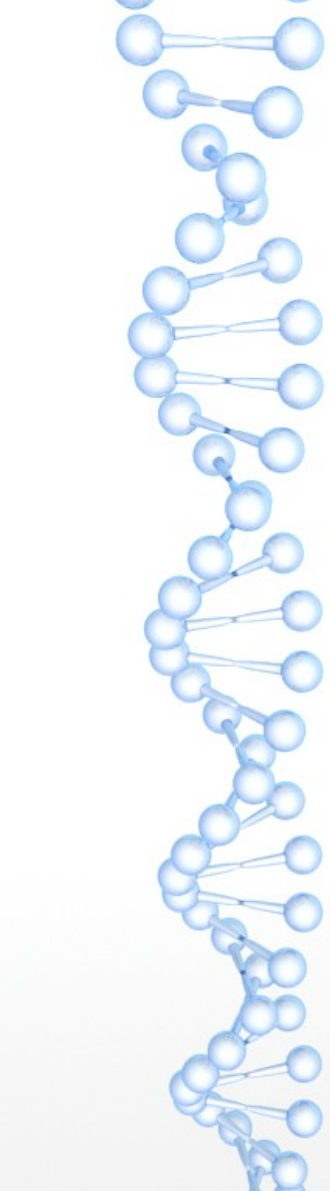
- Machine Learning algorithm ,“Model” – A mathematical expression that represents data in the context of a problem.

- Aim – Go from data to insight.



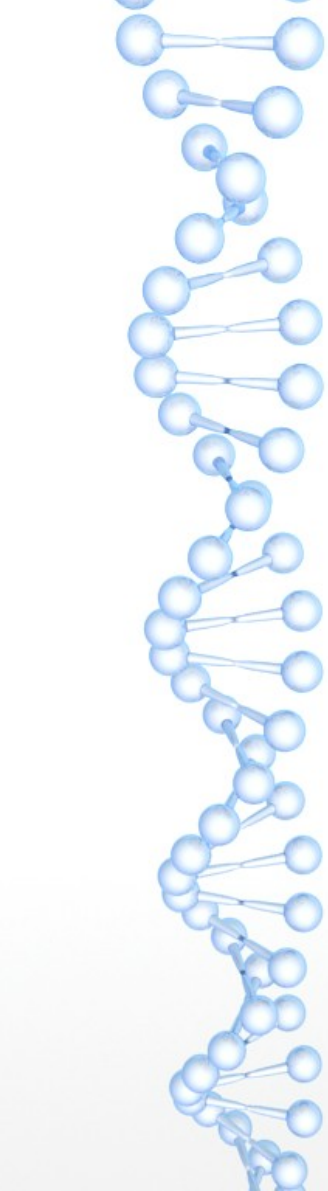
Methods

- Regression
- Classification
- Clustering
- Ensemble Methods
- Neural Networks & Deep Learning



Categories/Techniques

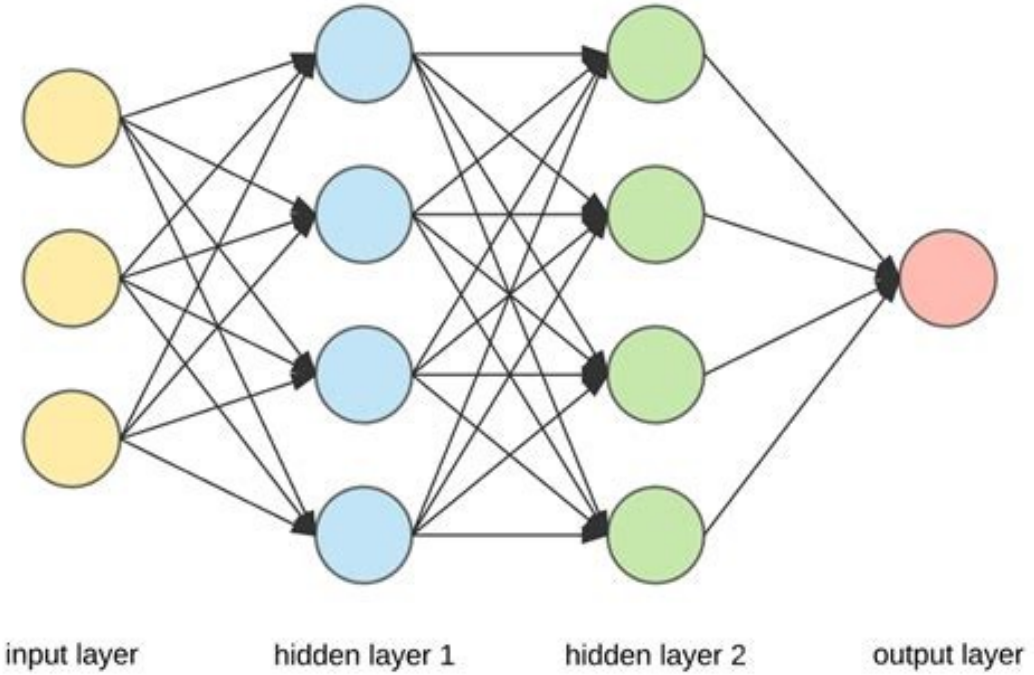
- Supervised – Have a piece of data that we want to explain or predict
- Unsupervised – Relate and group data points without the use of a target variable to predict.

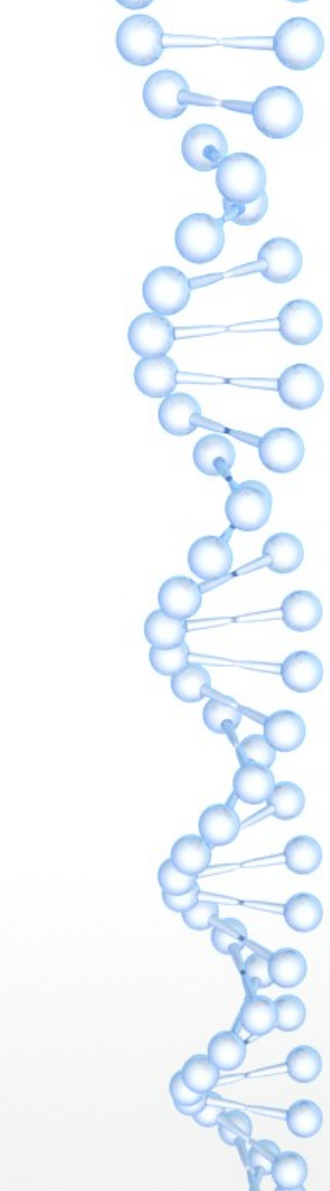


Deep Learning/Neural Networks.

- In contrast to linear models (Linear and Logistic regressions), Neural Networks aim at capturing non-linear patterns in data by adding layers of parameters to the model.

- Mimic Human human brain through a combination of data inputs, weights and bias.

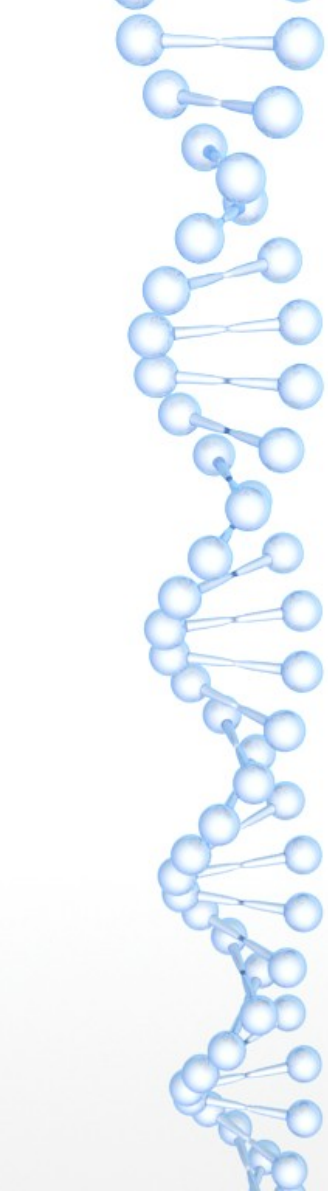

input layer
hidden layer 1
hidden layer 2
output layer



How do DNN work ?

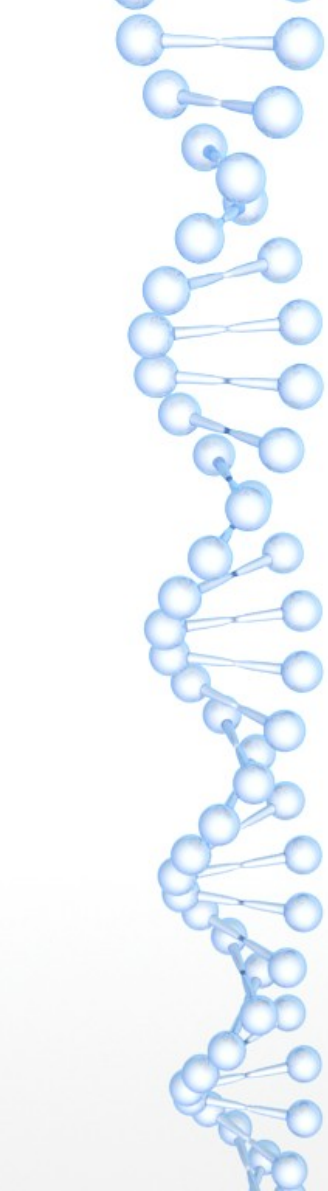
- Multiple layers of interconnected nodes, each building upon previous layers to refine and optimize the prediction or categorization (Forward propagation)

- Back propagation – Calculate errors in predictions and adjust the weights and biases by moving backwards through the layers in an effort to train the model.



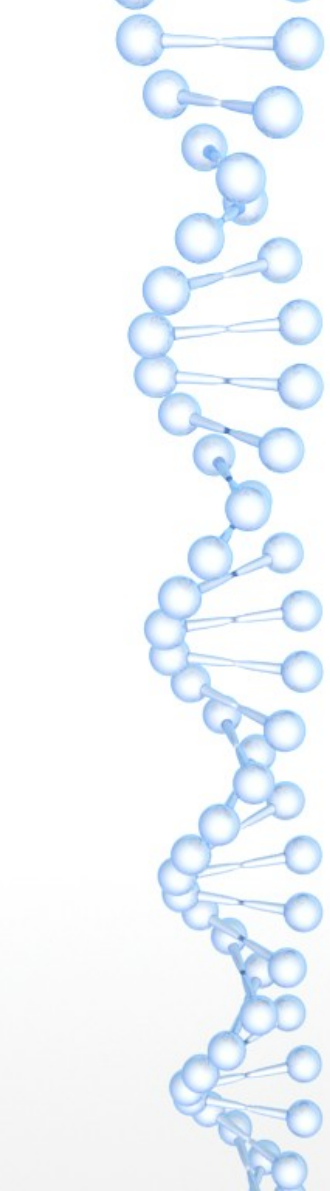
Weights & Biases

- Learnable parameters of a ML model
- Weights – Control signal; Decide how much influence the input will have on the output
- Biases – Constants; Always have a value of 1 ; Guarantees that even when all inputs are zeros there still will be an activation in the neuron.



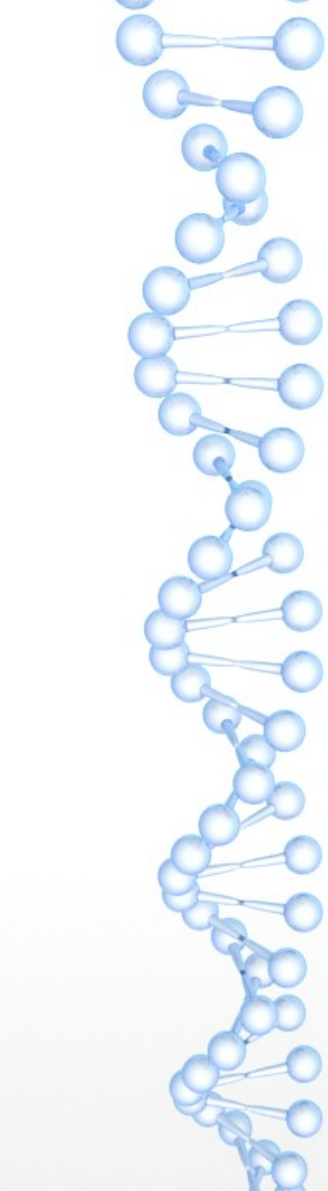
Activation Function

- Normalizes the input and produces an output which is then passed forward into the subsequent layer
- Add non-linearity to the output which enables neural networks to solve non-linear problems.
- **N/B: A NN without an activation function is just a linear regression.**
- Examples: Linear, Sigmoid, Tanh, ReLU.



Accuracy & Loss

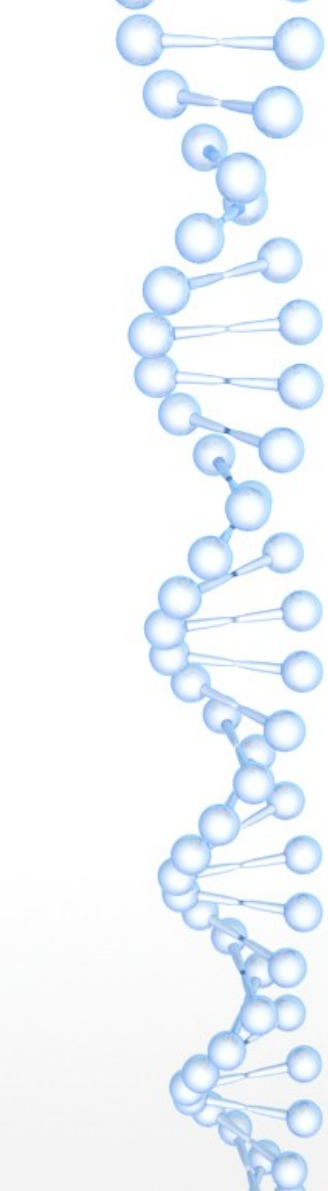
- Accuracy – Count of predictions where the predicted value is equal to the true value.(%) .
- Loss – Summation of the errors made for each sample in training or validation sets.
- Loss function / Cost function – Used for finding the best parameter values for the model (Weights) .
- Examples: Log loss, Cross-entropy loss, mean squared error and Likelihood loss.



Hyper parameters

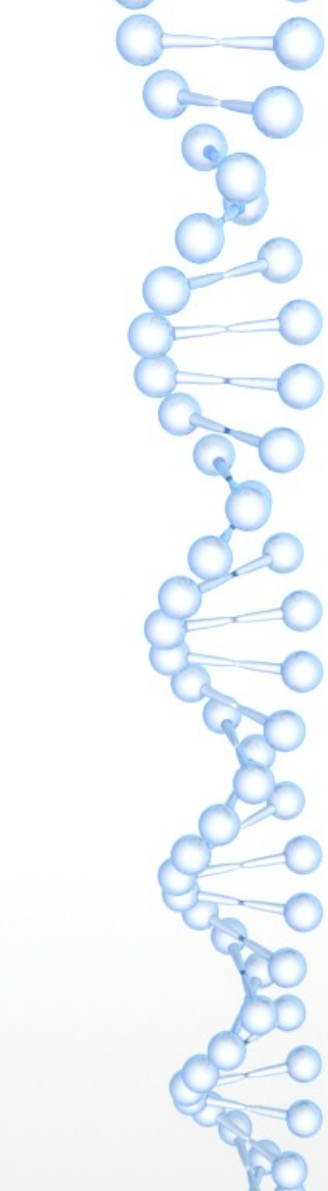
- A configuration external to the model whose value cannot be estimated from the data.

- Examples: No. of hidden layers, learning rate, activation function, No. of epochs.



Frameworks

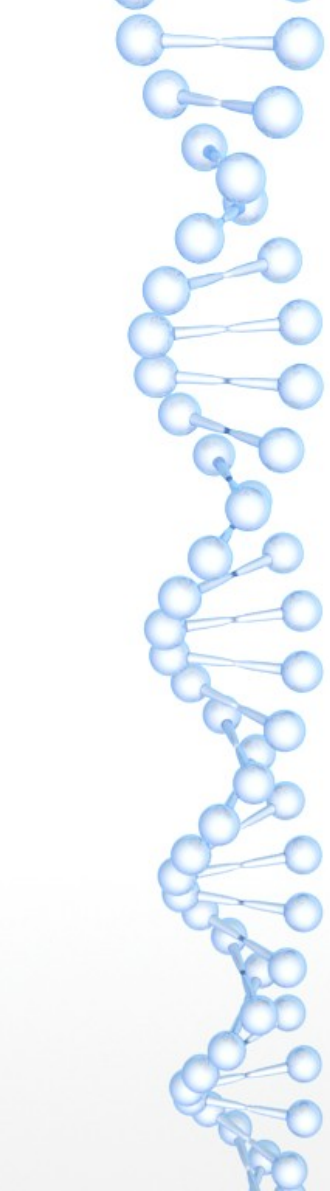
- Torch7
- Theano
- Caffe
- TensorFlow



TensorFlow

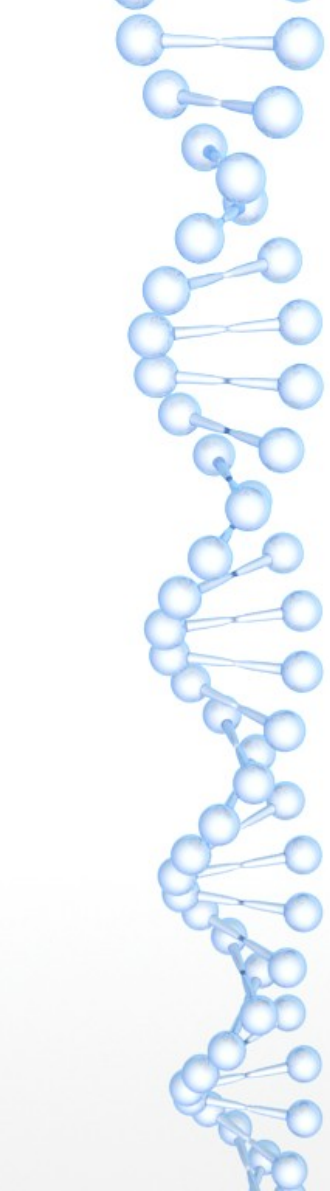
Pros

- Use declarative programming paradigm
- Allows for symbolic definitions
- Faster compilation times of the symbolic model
- Tensorboard for visualization
- Parralelization over multiple processors (CPU or GPU)
- Wrappers – Keras ;; Use R for the front-end
- Model deployments.



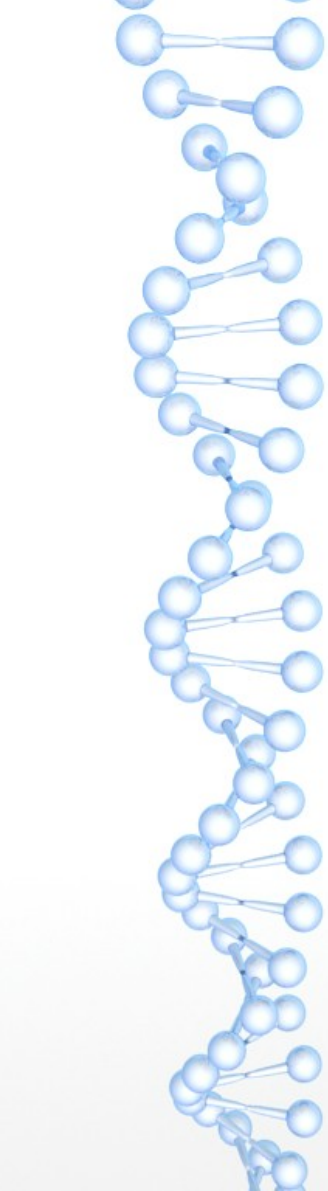
Deep learning in Computational Biology.

- Protein structure prediction – Earliest adopters of neural networks since 1980’s .
- A 50 – Layer deep learning model that improved contact map predictions on the CASP8 dataset by 10% .
- 2012 MERCK molecular activity challenge – A deep model won the competition



...

- Predicting various regulatory effects directly from DNA sequences.
- Predicted sequence specificities of DNA and RNA binding proteins (DeepBind)
- DeepSEA – Infer regulatory sequence code from chromatin profiling data.



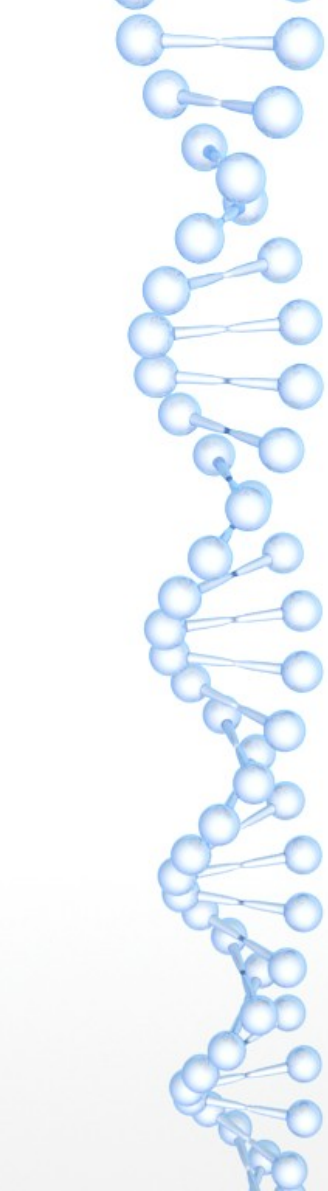
DL Challenges in Computational Biology

- Deep learning requires a dataset with many samples to work.
- Preparing biological data to fit in a deep learning model.

Solution:

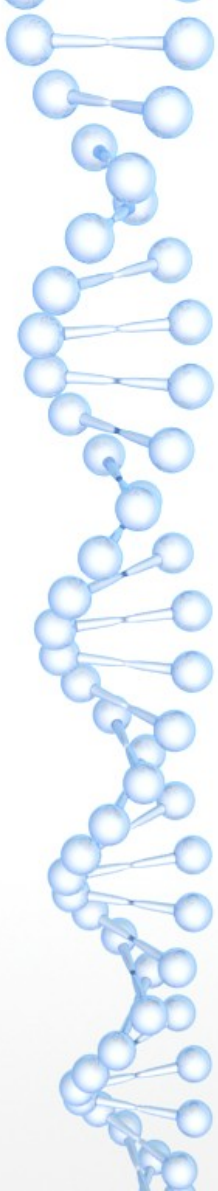
- Creating a scenario where the number of samples far exceeds the number of variables.

N/B : More workable solutions are still required here.



Conclusion

- TensorFlow is a powerful Deep learning framework.
- Pytorch a successor of Torch7 and Caffe is an equally powerful which offers the same functionality as TensorFlow.
- Whereas the whole concept of deep learning is the same for both libraries, Pytorch allows for a faster prototyping of a model as opposed to its counterpart.



THE END …