Affinity Prediction of Spinocerebellar Ataxia using Protein-Ligand and Protein-Protein Interactions with Functional Deep Learning

P. R. Asha, M. S. Vijaya

Abstract: ***Drug discovery of incomparable hereditary disorder like spinocerebellar ataxia is confronted and an enforce task in biomedical study. There are number of paths available for affinity prediction through scoring functions and ideals in the catalog. Nevertheless there is a need for artistic access in portraying the affinity of spinocerebellar ataxia which will facilitate enhanced prediction for drug discovery. This research work portrays the significance of docking for protein-ligand interaction and protein-protein interaction with modeling through deep learning. Deep Neural Networks is utilized in predicting binding affinity with 3d protein structures and ligand. Predictive models have been built with features related to for protein-ligand interaction and protein-protein interaction. In the first case, 17 protein structures and 18 ligands were used. Each protein structure is docked with ligand to get essential features like energy calculations, properties of protein and ligand for predicting binding affinity. In the next case, repeat mutation is induced manually with 17 protein structures and docked with 18 ligands. To train the model, well-defined descriptors are squeezed from the docked complex. Third case employs protein-protein interaction of total of 626 protein structures and the complexes attained from the protein-protein interaction are 313. Features like energy calculations, physio-chemical properties and interfacial and non-interfacial properties are extracted for learning this model. Deep learning has the property of representation learning from the user defined features, which helps in accurate prediction of binding affinity. The predictive models are developed with functional deep neural network and their performances are compared with sequential deep neural network. Functional deep neural network have more flexibility to define layers, complements sequential deep neural network which results in improved performance.***

Index Terms: ***Binding affinity, Deep Neural Network, Docking, Functional Deep Neural Network, Optimizers, Prediction, Protein Structure, Repeat Mutation***

I. INTRODUCTION

Affinity prediction of unusual genetic disorder is a crucial confront in biomedical research. Spinocerebellar Ataxia (SCA) is a rare genetic disorder characterized by degenerative development in the part of a brain correlated to the movement control and occasionally in spinal cord. SCA habitually have different types of symptoms for each type of disease and it has 36 types. People with SCA usually experience problem with coordination, uncoordinated walk, Abnormal speech, vision problems, poor hand-eye movement. SCA occurs by a change in a gene called mutation and most of the spinocerebellar ataxia occurs due to repeat mutation. Six types of spinocerebellar ataxia which occur through repeats mutation are SCA1, SCA2, SCA3, SCA6, SCA8 and SCA10. Trinucleotide repeat disorders is a collection of genetic disorders caused by trinucleotide repeat enlargement, a sort of mutation wherever trinucleotide repeats in bound genes or introns exceed the conventional, stable threshold, that differs per factor. The mutation may be a set of unstable microsatellite repeats that occur throughout all genomic sequences. If the repeat is gift in an exceedingly healthy factor, a dynamic mutation might increase the repeat count and lead to a defective factor. If the repeat is gift in associate degree deoxyribonucleic acid will cause toxic effects by forming spherical clusters known as polymer foci in cell nuclei. Trinucleotide repeats area unit typically classified as insertion mutations. Functional deep neural network allows defining multiple input or output models as well as models that share layers and also it allows in defining ad hoc acyclic network graphs. Sequential deep neural network is a linear stack of layers and it is limited in model topology whereas functional deep neural network is useful in creating complex models, such as multi-input/multi-output models, directed acyclic graphs, and models with shared layers. The functional deep neural network uses the same layers as the sequential deep neural network, but provides more flexibility in putting them together. In the functional deep neural network the layers are defined first, create the model, compile it, and fit it. Evaluation and prediction are essentially the same as in a sequential deep neural network. The layers are connected in pairwise by specifying where the input comes from when defining each new layer. A bracket notation is used, such that after the layer is created, the layer from which the input to the current layer comes from is specified. After creating the layers the models are defined and model defining is same as sequential deep neural network where training and testing takes place. Multi-task deep learning networks were introduced [Dahl, 2012] to predict the biomolecular target of 1 drug. Deep learning was accustomed notice off-target and cyanogenetic effects of environmental chemicals in nutrients, home product and medicines [Hochreiter's, 2014]. Deep neural network has recently started developing its applications in drug discovery, because of handiness of enormous chemical compounds and biological knowledge. Finding patterns between structures and activity, or structure and the other property, is referred to as a quantitative structure-activity analysis (QSAR). Single layer neural network has been utilized in QSAR, because of great deal of knowledge multilayer feed forward neural

Revised Manuscript Received on June 15, 2019

P. R. Asha, Department of Computer Science, Krishnammal College for Women, Coimbatore, India.

M. S. Vijaya, Department of Computer Science, Krishnammal College for Women, Coimbatore, India.

networks ar used for bioactivity predictions. a number of the learning's in compound property and activity prediction are: DNNs will handle thousands of descriptors while not the necessity of feature selection; (ii) dropout will avoid the disreputable overfitting drawback two-faced by a standard ANN; (iii) hyper-parameter (number of layers, range of nodes per layer, form of activation functions, etc.) optimisation will maximize the DNN performance; (iv) multitask DNN models perform higher than single-task models. DNN not only learns the hand-crafted features but also extracts the features by self while learning, and builds the accurate model.

II. LITERATURE REVIEW

Deep learning is developing in drug discovery process but very few contributions have been done with DNNs in the field of bioinformatics and drug discovery. Youjun Xu et al., [1] proposed a model for drug induced liver injury in deep learning. In this work DILI prediction models were developed using DL architectures. Model trained on 475 drugs predicted an external validation set of 198 drugs with an accuracy of 86.9%, sensitivity of 82.5%, specificity of 92.9%, and area under the curve of 0.955, which is better than the performance of previously described DILI prediction models. Authors used NCTR, Liew, Xu and combined datasets for both DILI positives and negatives. Datasets are passed to deep learning and the model is built for each and every dataset. The performance of DL-NCTR model with 190 drugs performs well with accuracy of 80.5% and with sensitivity 70.3%, specificity 88.2%. They performed internal and external cross validation for each and every dataset. DL-view model produce accuracy of 70% and 70% in sensitivity and 70% in specificity. DL- combined model performs best with accuracy of 88.9% and 89.9%, 87% of sensitivity and specificity accordingly. The combined data was constructed because greene, Xu and NCTR were all drug datasets and they could be combined. While combining these three datasets the dataset becomes large, that would be beneficial in determining the advantage of deep learning. UGRNN encoding approach and with large datasets authors have developed the prediction of DILI drugs and small compounds. The DL-combined model performs best with highest accuracy. Bharath Ramsundar et al., [2] scrutinizes the use of massively multitask networks for drug discovery. Dataset was created by gathering large amount of data from public sources, nearly 40 million measurements across 200 biological targets. The purpose of this work is, multitask network performs better than single-task methods. The predictive power of multitask increases as additional tasks and data are added. Models were trained on 259 datasets gathered from public sources. These datasets were divided into four groups PCBA, MUV, DUD-E, and Tox21. PCBA group contained 128 experiments in the PubChem BioAssay database. MUV group contained 17 challenging datasets specifically designed to avoid common pitfalls in virtual screening. The DUD-E group contained 102 datasets that were designed for the evaluation of methods to predict interactions between proteins and small molecules. The Tox21 datasets were used in the recent Tox21 Data Challenge and contained experimental data for 12 targets relevant to drug toxicity prediction. Extended connectivity fingerprints were generated to featurize each molecule. Models that are built using multitasking network are logistic regression, random forest, single task neural network, max, pyramidal, one-hidden layer multitask neural network and pyramidal multitask neural network. The efficient of multitasking network is availability of relevant data. There is a critical need for greater amount of data.Babak Alipanahi [3] presented that the prediction of sequence specificities of DNA and RNA binding proteins by deep learning. DeepBind approach has been used and built standalone software tool that is fully automatic and handles millions of sequences per experiment. In DeepBind approaches there are following challenges like (i) It can be applied to both microarray and sequencing data; (ii) it can learn from millions of sequences through parallel implementation on a graphics processing unit (GPU); (iii) it generalizes well across technologies, even without correcting for technology-specific biases; (iv) it can tolerate a moderate degree of noise and mislabeled training data; and (v) it can train predictive models fully automatically, alleviating the need for careful and time-consuming hand-tuning. It uses a set of sequences and, for each sequence, an experimentally determined binding score. Sequences can have varying lengths (14–101 in our experiments), and binding scores can be real-valued measurements or binary class labels. Deep learning outperforms other state-of-art methods, even when training on in vitro data and testing on in vivo data.The background study evidently case that there is a decisive require for affinity prediction, because number of disease amplify swiftly. Since there is no treatment for rare genetic disorders, the system facilitates the improvement of drug to guide clinicians in efficient drug detection of SCA with repeat mutations. The research work aims at predicting the binding affinity of SCA using various datasets like protein-ligand, protein-ligand with mutation induced in protein structures and protein-protein interaction, which is carried out with functional API.

III. METHODOLOGY

This work employs deep learning based functional model to develop predictive models for predicting binding affinity. The framework of the functional model shown in Fig. 1 is capable of creating instance of layers, connecting the layers and defining model, which specifies layers to act as input layer and output layer. In functional deep neural network standalone input layer is created, which specifies the shape of data, where as in sequential deep neural network, it is limited in creating models which do not share layers or possess multiple inputs or outputs. Functional model delineates model effortlessly where layers unite to more than one layers. In order to characterize the benefits brought about by functional deep neural network, it is compared to sequential deep neural network. The input layer acquires a shape argument that is an ordered pair that demonstrates the dimensionality of the input data. Deep learning neural network with functional model provide a better prediction through hyperparameters optimization. The work is divided into three phases namely creation of dataset, model building and performance evaluation.

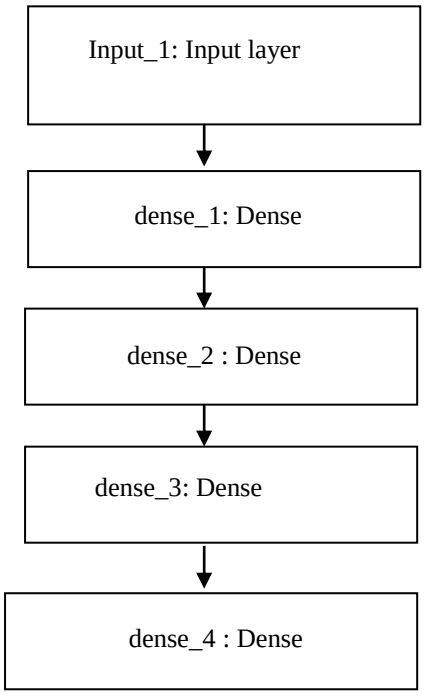

Fig 1. Deep Learning based Functional Model Framework

A. Datasets

Three datasets have been developed in this research work for experimentation on functional deep neural network. Protein-Ligand docking (PLD) is focused to develop the first dataset, whereas in the second case the Protein Mutated-Ligand docking (PMLD) is taken into account to define the features. Protein-Protein Interaction (PPID) is adopted to fetch the features for third dataset. The procedures for creating these datasets are presented below.

PLD Dataset

3d structures are compiled from protein data bank and the dataset is equipped from the docked complex of 17 structures and 18 ligands. Structures for repeat mutation are gathered from gene cards and the ligand information is collected from literatures. Flexible docking is performed, where ligand rotates the protein and for every rotation, energy is calculated. Optimum pose is selected to get docked complex and to organize the dataset each protein structure is docked with ligand. Indispensable features are extracted from the docked complex.PLD constitutes features like energy calculations and physical properties of ligand, protein are extracted from autodock, autodock vina and pymol. Binding affinity is measured as dependant variable and other feature values are considered as predictor variables. Binding affinity is evaluated using

$$[R]\,[L]\,k1 = [RL]\,k\text{-}1$$
$$k1/k{-}1 = [RL]/[R][L]$$
$$\text{Binding Affinity} = k1/k{-}1$$
$$Kd = k{-}1/k1$$

Kd is known as binding affinity.

k1 is termed as association constant for forward reaction and k-1 is defined as rate constant for backward reaction. The PLD dataset consists of 28 attributes and 307 instances.

PMLD Dataset

In the second case, protein structures are unruffled from gene cards for six types of spinocerebellar ataxia which ensue owing to repeat mutation. A total of 17 structures are collected and repeat mutations are induced manually. Mutated protein structures are docked with ligand and features are elicited from the docked complex, the dataset is encoded as PMLD (Protein Mutated-Ligand Dataset). Features like energy based descriptors, cyscore, rfscore, sequence descriptors and autodock vina scores are elicited from the docked complex. PMLD compasses of 306 instances and 509 attributes and the features are energy based descriptors consists of 7 descriptors namely binding energy, inhibition constant, intermolecular energy, desolvation energy, electrostatic energy, total internal energy and torsional energy. Sequence descriptors consist of 1097 descriptors namely amino acid composition, autocorrelation, CTD, Quasi-sequence-order descriptors, Pseudo amino acid composition and Profile-based descriptors. Cyscore posse's five descriptors to be precise Hydrophobic free energy, cyscore, van der waals interaction energy, hydrogen-bond interaction and ligand's conformational entropy. Rfscore consist of thirty six values and each feature will comprise the number of occurrences of a particular protein-ligand atom type pair interacting within a certain distance range. Autodock vina scores has four descriptors namely ΔG_{gauss}, $\Delta G_{repulsion}$, $\Delta G_{hydrophobic}$ and ΔG_{Hbond}.

PPID Datasets

For the third dataset, protein structures of the SCA genes that are pretentious by repeat mutations are serene from PDB database and the interaction information about the proteins is composed from genecards. Rigid docking is performed for protein-protein interaction and the molecular objects are handled as rigid. Hotspot is acknowledged by transitory the protein structure into convolutional neural network. Threshold value is preset and the amino acid in that complex which procures greater than 0.5 is treated for hotspots. Protein interaction produces many clusters and the finest cluster is preferred by the least minimum energy and the dataset is hard-coded as PPID (Protein-Protein Interaction Dataset).PPID has 313 instances and 56 attributes and features are distilled from the interacted complex and the features are haddock score, cluster size, RMSD, desolvation energy, van der waals energy, electrostatic energy, z-score, binding affinity, physio-chemical properties, interfacial contacts and NIS properties. In account to these features DNN extracts self-learned features, which is more reliable for predicting affinity prediction. Profile of the above datasets is listed in Table 1. Each dataset is split into training and test sets for building the predictive models and evaluating their performances using a standard 10- fold cross-validation technique.

Table 1. Profile of Datasets

Datasets	Concept	Summary of Protein Structures and Ligand	Summary of Features	No. of Instances
PLD	Protein-Ligand Docking	17 structrues and 18 Ligands	28	307
PMLD	Protein Mutated-Ligand	17 structures and 18 Ligands	509	306
PPID	Protein-Protein Interaction	626 structures	56	313

B. Model Building

Binding affinity prediction is crucial for drug identification of rare genetic disorder and hence it is proposed in this research work to build efficient model for predicting binding affinity. The problem of the predicting binding affinity is modeled as regression task and solved using functional deep neural network. For this purpose three datasets described above have been used to train the models. Feature values are normalized using Min-Max normalization. Various binding affinity predictive models are learned with three optimizers namely Adam, RMSprop and Nadam. The system planning for affinity prediction model is shown in Fig. 2.Functional deep neural network is built by creating the instances of layer and connecting them to each other. The input is specified by giving the number of dimensions and that input layer is shared with hidden layer. Deep neural network extracts intellectual features by itself using its kernel and learns high level features, which is more reliable for building the model. It is essential to specify input and output layer for building the functional deep neural network and it is experimented with Adam optimizer. The model is optimized by the hyper parameters such as epochs, dropout rate, optimizers and learning rate.The proposed model uses a dense layer for prediction and 30 hidden layers with one output layer. In this work, many dense layers are connected with one another to perform functional models. Input dimension is 55 and for this input dimension hidden layers can be changed accordingly and the output layer is a single layer. A Relu activation function is used on the output layer to allow the network to learn and output the distribution over the possible output values. The network uses the mean squared loss function while training, which is suitable for prediction problems. Functional deep learning neural network is implemented with three different optimizers. Adam optimizer is conservatory of stochastic gradient optimization algorithm and Adam optimization algorithm computationally renovates network weights iterative based on training data. The parameter values used are learning_rate=0.01, beta1=0.9, beta2=0.999, epsilon=1e-08. Instead of adapting the parameter learning rates based on the average first moment, Adam also makes use of the average of the second moments of the gradients. Specifically, the algorithm appraises an epidemic poignant average of the gradient and the squared gradient, and the parameters beta1 and beta2 control the decay rates of these moving averages. The prediction metrics are expressed for each training epoch to furnish an idea of the skill of the model in addition to the loss.Learning rate is the parameter that implies the optimizer conversely so much to contrive the weights within the direction opposite of the gradient for a mini-batch. Weight changes are often massive that the optimizer overshoots the minimum and makes the loss worse. Different values of 0, 0.1, 0.01 and 0.001 were tested to search out the one that offers the simplest loss while not sacrificing speed of coaching. When training with a smaller learning rate, the error rate will be low, whereas when it is performed with higher learning rate the error rate converges. Hence 0.01 was mounted as the learning rate.

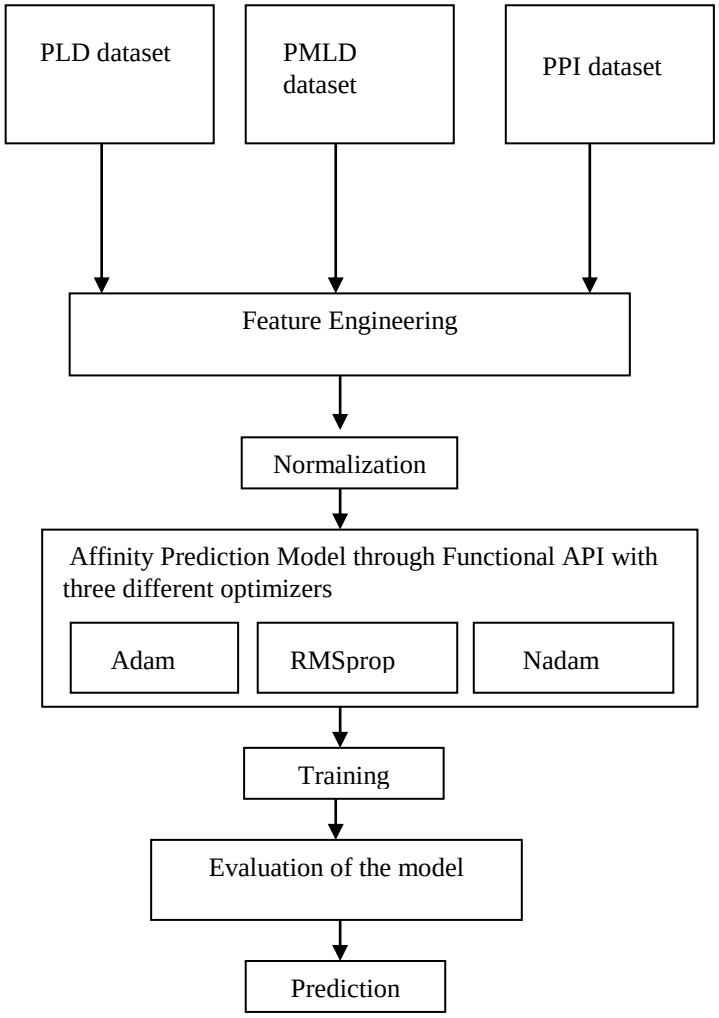

Fig. 2 System Planning of Binding Affinity Predictive Models

Regularization is a great approach to restrain overfitting the training data. Dropout is a regularization parameter that randomly skips neurons during training, forcing others in the layer to pick up the slack. The contribution of the detached neurons to the activation of downstream neurons is temporally removed on the forward pass and any weight updates are not applied to the neuron on the backward pass. Dropout as a hyperparameter is implemented by randomly selecting nodes to be dropped-out with a given probability in each weight update cycle. The model was tested with different dropout percentages varying from 20- 50% and the results were recordedThree independent predictive models are built with functional model of deep learning with three optimizers. Each model is trained for all the three datasets and the performance measures evaluated for these models to predict affinity are explained variance score, mean squared error, root mean squared error, R2 score, mean absolute error and median absolute error. The performance of functional deep neural network is compared with sequential deep neural network.

IV. EXPERIMENT AND RESULTS

Binding affinity predictive model trained with PLD dataset using functional deep neural network is experimented with three optimizers like Adam, RMSprop and Nadam. Each optimizer possesses varying number of epochs and dropout values. Among the three optimizers, each optimizer attains the superlative performance at 500 epochs and at 0.3 dropout value. Adam optimizer outperforms the other two optimizers, which attains best performance of 0.92. Metrics of all the three optimizers are listed in Table 2 and Fig. 3 is shown for varying dropouts and metrics like explained variance score and mean squared error, which is considered most significant. Three independent experiments have been carried out with functional model of deep neural network. Functional deep neural network is implemented in keras platform as front end and tensorflow as back end with deep neural network. Three affinity predictive models using functional deep neural network are built with three optimizers namely Adam, RMSprop and Nadam. Adam optimizer with different parameters like learning rate, epoch, dropout and activation functions is used with 55 input dimensions. Results show that functional deep neural network with Adam optimizer outperforms contending sequential deep neural network.

Table 2. Results of Binding Affinity Predictive Model with PLD Dataset using Functional Deep Neural Network

Dropout	Epochs	Optimizer	Explained variance score	R2 score	Mean squared error	Root mean squared error	Mean absolute error	Median absolute error
0.3	500	Adam	0.92	0.92	0.14	0.32	0.12	0.4
		RMSprop	0.90	0.90	0.15	0.34	0.15	0.6
		Nadam	0.85	0.85	0.20	0.44	0.25	0.15

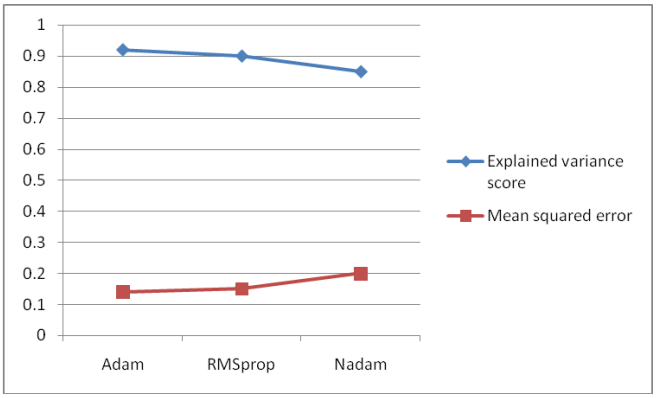

Fig 3. Results of Binding Affinity Predictive Model with PLD Dataset using Functional Deep Neural Network

To boot the performance of the binding affinity predictive model, it is trained with PMLD dataset using functional deep neural network. The model is implemented with three optimizers like Adam, RMSprop and Nadam. Each optimizer has varying number of epochs and dropout values. Among the three optimizers, each optimizer gains the exceptional performance at 500 epochs and at 0.3 dropout value. In the second model too Adam optimizer outperforms the other two optimizers, which attains best performance of 0.90. Metrics of PMLD dataset with three optimizers are listed in Table 3 and Fig. 4 is shown for metrics like explained variance score and mean squared error.

Table 3. Results of Binding Affinity Predictive Model with PMLD Dataset using Functional Deep Neural Network

Dropout	Epochs	Optimizer	Explained variance score	R2 score	Mean squared error	Root mean squared error	Mean absolute error	Median absolute error

0.3	500	Adam	0.90	0.90	0.15	0.34	0.15	0.6
		RMSprop	0.85	0.85	0.20	0.44	0.25	0.15
		Nadam	0.75	0.75	0.30	0.59	0.39	0.27

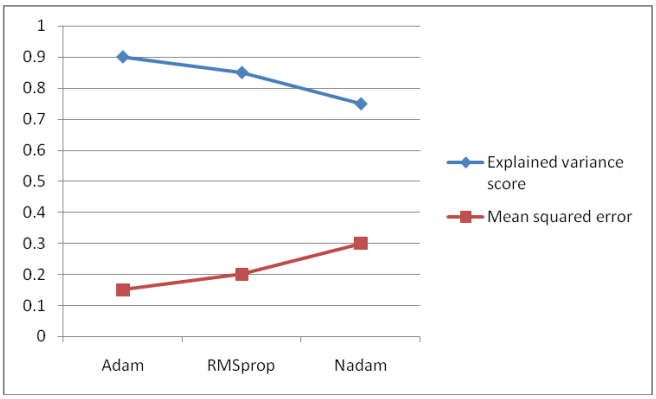

Fig 4. Results of Binding Affinity Predictive Model with PMLD Dataset using Functional Deep Neural Network

Binding affinity predictive model trained with protein-protein dataset and using functional deep neural network is performed with three optimizers like Adam, RMSprop and Nadam. Every optimizer has varying number of epochs and dropout values. Amongst the three optimizers, each optimizer gains the excellent performance at 500 epochs and at 0.3 dropout value. Adam optimizer produces enhanced recital than other two optimizers, which attains best performance of 0.95. Metrics of PPI dataset with three optimizers are listed in Table 4 and Fig. 5 is exposed for metrics like explained variance score and mean squared error.

Table 4. Results of Binding Affinity Predictive model with PPI Dataset with Functional Deep Neural Network

Dropout	Epochs	Optimizer	Explained variance score	R2 score	Mean squared error	Root mean squared error	Mean absolute error	Median absolute error
0.3	500	Adam	0.97	0.97	0.1	0.32	0.12	0.4
		RMSprop	0.92	0.92	0.14	0.32	0.12	0.4
		Nadam	0.90	0.90	0.15	0.34	0.15	0.6

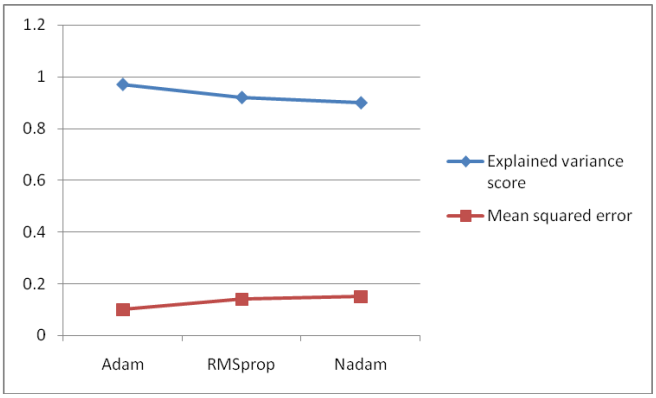

Fig 5.Results of Binding Affinity Predictive Model with PPI Dataset using Functional Deep Neural Network

Experimental results illustrate that the Adam optimizer with functional deep neural network outperforms rival optimizers. The performance of all the three models is compared with the sequential deep neural network of Adam optimizer based on deep neural network. Sequential deep neural network has the same number of epochs and dropout values and both models achieve the best performance at same epoch and dropout value like 500 and 0.3 respectively. The explained_variance score and mean_squared error of both binding affinity predictive models with sequential and functional deep neural network with adam optimizer of varying dropouts models are given in Table 5 and Fig. 6 is shown for sequential deep neural network. Functional deep neural network is shown in Fig. 7 and it is obvious that the functional deep neural network achieved comparable performance than the sequential deep neural network. The PPI dataset with both sequential and functional deep neural network is more accurate than the PLD dataset and PMLD dataset.

Table 5. Results of Binding Affinity Predictive Model with Sequential and Functional Deep Neural Network with Adam Optimizer

Dataset	Model name	Explained variance score	Mean squared error

PLD	**Sequential**	**0.82**	**0.20**
PMLD		**0.86**	**0.2**
PPI		**0.95**	**0.1**
PLD	**Functional**	**0.86**	**0.2**
PMLD		**0.90**	**0.15**
PPI		**0.97**	**0.1**

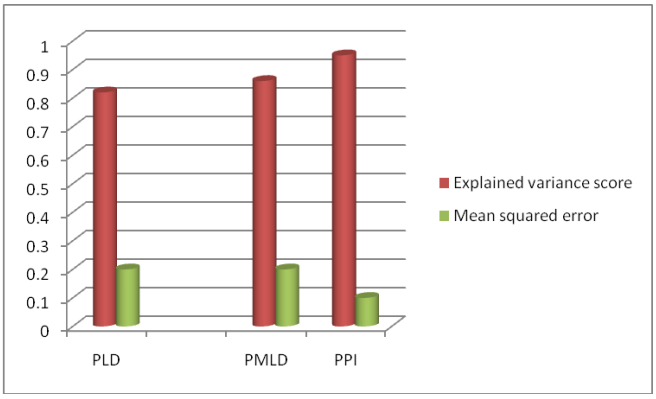

Fig 6. Results of Binding Affinity Predictive Models with Sequential Deep Neural Network for Three Datasets

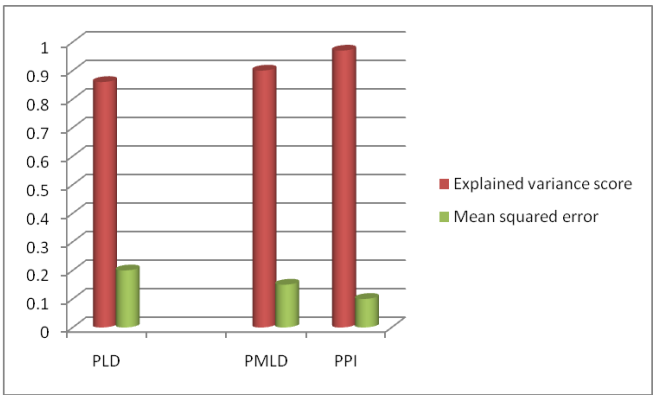

Fig 7. Results of Binding Affinity Predictive Models with Functional Deep Neural Network for Three Datasets

V. FINDINGS

From the above experiments, it is obviously inferred that functional deep neural network can predict the binding affinity with better results. The suitable choice of hyperparameters improves the prediction performance of the models. Functional deep neural network learns complex, hidden and high level features from the hand crafted features like energy calculations, scores from haddock and physio-chemical properties through several levels of hidden units. The high recognition rate obtained by functional deep neural network proves that the models are effective in predicting binding affinity. The framework of functional deep neural network enabled feature learning and regression within the deep architecture which helps to reduce the number of tasks in building models through shallow learning and to predict binding affinity accurately. It is inveterate that the model built using functional model with Adam optimizer conquers enhanced results for affinity prediction from protein-protein interaction complexes as providential outcomes are achieved. Explained_variance score is very prominent and the Mean squared error is minimized so the reliability of the system is improved. Functional deep learning architecture implemented using tensorflow creates instances of layer and connects the layer, which helps in solving the complex models. It is also evident that this learning technique, with Adam optimizer is appropriate in predicting binding affinity of spinocerebellar ataxia and also for any rare genetic disorder. This research work proves that model built using functional deep neural network outperforms the model built with sequential deep neural network.

VI. CONCLUSION

In this paper, the influence of protein-protein interaction in predicting binding affinity is illustrated through functional deep neural network. Functional deep neural network is trained with various optimization functions and regulated hyper-parameters. The effectiveness of protein-protein interaction in binding affinity prediction is proved by comparing with models based on protein-ligand docking and protein mutated-ligand docking. The crucial task of self feature representation in DNN facilitates straightforward modeling of binding affinity. The experimental results confirmed that the functional model outperforms sequential deep neural network and demonstrates that the functional model can make more effective affinity prediction. Result analysis provides a baseline for future research, and it is expected that it can give a better result when using combined network configurations. In future, the effects of adding depth to the deep neural network and varying activation function can be explored.

REFERENCES

1. *Youjun Xu, Ziwei Dai, Fangjin Chen, Shuaishi Gao, Jianfeng Pei, and Luhua Lai, "Deep Learning for Drug-Induced Liver Injury", Journal of Chemical Information and Modeling 2015 55 (10), 2085-2093*
2. *Bharath Ramsundar, Steven Keames, Patrick Riley, Dale Webster, David Konerding, Vijay Pande, "Massively Multitask Networks for Drug Discovery", Statistics, Feburary 6, 2015*
3. *Babak Alipanahi, Andrew Delong, Matthew T Weirauch & Brendan J Frey, Predicting the sequence specificities of DNA- and RNA-binding proteins by deep learning, Nature Biotechnology, July 27, 2015*
4. *HongmingChen, OlaEngkvist, YinhaiWang, MarcusOlivecrona, ThomasBlaschke, The rise of deep learning in drug discovery, vol 23, issue 6, pg 1241-1250, june 2018.*
5. *Izhar Wallach, Michael Dzamba, Abraham Heifets, AtomNet: A Deep Convolutional Neural Network for Bioactivity Prediction in Structure-based Drug Discovery, cornell university library, oct 10, 2015*
6. *Babak Alipanahi, Andrew Delong, Matthew T Weirauch, and Brendan J Frey. Predicting the sequence specificities of dna-and rna-binding proteins by deep learning, Nature biotechnology,2015.*
7. *Alessandro Lusci, Gianluca Pollastri, and Pierre Baldi, Deep Architectures and Deep Learning in Chemoinformatics: The Prediction of Aqueous Solubility for Drug-Like Molecules, journal of chemical information and modeling 2013 53(7), 1563-1575*

8. *Alexander Aliper, Sergey Plis, Artem Artemov, Alvaro Ulloa, Polina Mamoshina, and Alex Zhavoronkov, Deep learning applications for predicting pharmacological properties of drugs and drug repurposing using transcriptomic data, Molecular Pharmaceutics, 2016 13(7), 2524-2530*

AUTHORS PROFILE

P. R. Asha, Ph. D Research Scholar,Completed M.Sc[SE] in Bannari Amman Institute, M.phi in krishnammal college for women and currently pursuing my Ph.D in krishnammal college for women.Areas of Interest: Computational biology, data mining and statistics. I'm interested in finding new drugs for the disease which has no drug to cure the disease. Pursuing research work in predicting binding affinity for spinocerebellar ataxia. Four papers has been presented and published in conference proceedings. Paper published in **JARDCS** journal entitled "Deep Neural Networks for Affinity Prediction of Spinocerebellar Ataxia Using Protein Structures".

M. S. Vijaya, Associate ProfessorCompleted masters in PSG college of Technology and did Ph.D in Amirta university Coimbatore and currently working as associate professor in krishnammal college for women. She is an head of computer science department.Area of Specialization: Data Mining, Machine Learning, Support Vector Machine, Pattern Recognition. She has guided many M.phil research scholars and one Ph.d research scholar completed under her guidance. Many Ph.d research scholars has been submitted their thesis and has been waiting for viva voice.