International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 06 Issue: 07 | July - 2022 Impact Factor: 7.185 ISSN: 2582-3930

Chronic Disease Prediction Using Machine Learning

Tippana Chalapathi Rao[1], Injarapu Venkata Vara Sai Phaneendra[2], Andhavarapu Tejitha[3]

Jami Sai Surya[4], Maddi Keerthi Sudha[5]

[1] Sr. Assistant Professor, [2,3,4,5]Student

[1,2,3,4,5] Department of Computer Science Engineering

[1,2,3,4,5] Aditya Institute of Technology and Management, Tekkali, Andhra Pradesh

ABSTRACT

Technological development, including machine learning, has a huge impact on health through an effective analysis of various chronic diseases for more accurate diagnosis and successful treatment. In this digital world, most of the people are prone to diseases, due to lack of healthy food, proper sleep and daily exercise. It is very important to know if we are suffering from a disease, at an early stage rather than discovering it at a later stage. Hence disease prediction system plays an important role as it predicts the diseases based on system. The objective of this project is to predict the Pattern Classification by using Neural Network (NN)and Adam Optimization Algorithm (Adam). In this project we used machine learning for training the data sets. The data sets include the records of the patients who are infected. Here we are taken the datasets of the diseases that are Cancer, Liver, Diabetes. We introduce Adam, an algorithm for first-order gradient-based optimization of stochastic objective functions, based on adaptive estimates of lower-order moments. The method is straightforward to implement, is computationally efficient, has little memory requirements. The obtained results show that the proposed method performs better then or on par with other methods in terms of classification accuracy and sum squared errors and it predicts whether the person is infected or not.

Keywords: Adam Optimization Algorithm, Multilayer Perceptron Neural Networks, Pattern Classifications, UCI Datasets, Machine Learning.

INTRODUCTION

A major challenge faced by health care organizations, such as hospitals and medical Centre, is the provision of quality services at affordable costs. The quality service implies diagnosing patients properly and administering effective treatments. The available chronic disease database consists of both numerical and categorical data. Before further processing, cleaning and filtering are applied on these records in order to filter the irrelevant data from the database. The proposed system can determine an exact hidden knowledge, i.e, patterns and relationships.

The name Adam is gotten from versatile second estimation. Adam is a well-known calculation in the field of profound learning since it accomplishes great outcomes quick. It is suggested by some notable neural system calculation specialists. Streamlining is characterized as the determination of the best arrangement. Essential outcomes and numerical strategies in enhancement can be utilized to locate the perfect decision among numerous potential other options. Improvement strategies are of deterministic and arbitrary insight approaches. Deterministic techniques produce indistinct course of action if its starting qualities are equivalent with each other when dealing with a comparative issue. Nature gives a part of the beneficial ways to deal with deal with some mind-boggling issues or issues .

Calculations that are copying structures in nature/motivated from nature are called Nature Inspired Algorithms. Nature-persuaded estimations are novel, basic reasoning strategies and have been pulling in noteworthy thought for their extraordinary execution. Agent occasions of nature motivated calculations incorporate Swarm Intelligence (SI), Artificial Neural Networks (ANN) and Evolutionary Computing (EC), and so forth. These nature moving calculations are having significant job in taking care of numerous genuine complex issues by utilizing advancement procedures and models. Metaheuristic enhancement procedures have gotten extremely mainstream throughout the most recent two decades.

Optimization is defined as the selection of the best solution. Fundamental results and numerical methods in optimization can be used to find the ideal choice among many possible alternatives. Optimization methods are of deterministic and random intelligence approaches. Deterministic methodologies produce indistinguishable arrangement if its beginning values are equal with one another when taking care of a similar issue. Unlike deterministic approaches, gradient free stochastic approaches are mainly based on random walks

LITERATURE REVIEW

Adam is an adaptive learning rate optimization algorithm that's been designed specifically for training deep neural networks. First published in 2014, Adam was presented at a very prestigious conference for deep learning practitioners. ICLR 2015 The paper contained some very promising diagrams, showing huge performance gains in terms of speed of training. However, after a while people started noticing, that in some cases Adam actually finds worse solution than stochastic gradient descent. A lot of research has been done to address the problems of Adam.

The algorithms leverage the power of adaptive learning rates methods to find individual learning rates for each parameter. It also has advantages of Ad grad, which works really well in settings with sparse gradients, but struggles in non-convex optimization of neural networks, and RMSprop, which tackles to resolve some of the problems of Ad grad and works really well in on-line settings. Adam has been raising in popularity exponentially according to 'A Peek at Trends in Machine Learning' article from Andrej Karpathy.

In this post, I first introduce Adam algorithm as presented in the original paper, and then walk-through latest research around it that demonstrates some potential reasons why the algorithms works worse than classic SGD in some areas and provides several solutions, that narrow the gap between SGD and Adam.

ADAM OPTIMIZATION

Adam involves the following steps:

1. Compute the gradient and its element-wise square using the current parameters.
2. Update the exponential moving average of the 1st-order moment and the 2nd-ordermoment.
3. Compute an unbiased average of the 1st-order moment and 2nd-order moment.
4. Compute weight update: 1st-order moment unbiased average divided by the square rootof 2nd- order moment unbiased average (and scale by learning rate).
5. Apply update to the weights.

Advantages of using Adam on non-convex optimization issues:

- Implementation is straightforward and Effective in computing.
- Invariant of gradient diagonal rescale.
- Best suited for information or parameters-sized problems.

- Suitable for non-stationary targets.
- Suitable for very noisy/sparse gradient problems.
- Easy to implement and Quite computationally efficient
- Requires little memory space
- Good for non-stationary objectives
- Works well on problems with noisy or sparse gradients
- Works well with large data sets and large parameters

DATA COLLECTION

Diabetes Attributes: Pregnancies, Glucose, Blood Pressure, Skin Thickness, Insulin, BMI, Diabetes Pedigree Function, Age, Outcome.

Liver Tumor Attributes: Total Bilirubin, Alkaline Phosphatase, Alaina Aminotransferase, Total Proteins, Albumin, Albumin and Globulin ratio

Cancer Attributes: Cell Thickness**,** Cell size, Cell Shape, Bare Nuclei**,** Norm Nucleoli**,** Mitoses**.**

EXPERIMENTAL RESULT

DIABETES:

```
     Pregnancies  Glucose  BloodPressure  SkinThickness  Insulin   BMI  \
0              6      148             72             35        0  33.6
1              1       85             66             29        0  26.6
2              8      183             64              0        0  23.3
3              1       89             66             23       94  28.1
4              0      137             40             35      168  43.1
..           ...      ...            ...            ...      ...   ...
164           10      101             76             48      180  32.9
165            2      122             70             27        0  36.8
166            5      121             72             23      112  26.2
167            1      126             60              0        0  30.1
168            1       93             70             31        0  30.4

     DiabetesPedigreeFunction  Age
0                       0.627   50
1                       0.351   31
2                       0.672   32
3                       0.167   21
4                       2.288   33
..                        ...  ...
164                     0.171   63
165                     0.340   27
166                     0.245   30
167                     0.349   47
168                     0.315   23

[169 rows x 8 columns]
[0.7647059  0.64705884 0.7647059  0.52941179 0.41176471 0.52941179
 0.58823532 0.64705884 0.64705884 0.6875    ]
Diabetes Baseline: 62.17% (10.46%)
[1]
Positive
```

LIVER:

	Total_Bilirubin	Direct_Bilirubin	Alkaline_Phosphotase	Alamine_Aminotransferase	Aspartate_Aminotransferase	Total_Protiens	Albumin	Albumin_and_Globulin_Ratio
0	0.7	0.1	187	16	18	6.8	3.3	0.90
1	10.9	5.5	699	64	100	7.5	3.2	0.74
2	7.3	4.1	490	60	68	7.0	3.3	0.89
3	1.0	0.4	182	14	20	6.8	3.4	1.00
4	3.9	2.0	195	27	59	7.3	2.4	0.40
5	1.8	0.7	208	19	14	7.6	4.4	1.30
6	0.9	0.2	154	16	12	7.0	3.5	1.00
7	0.9	0.3	202	14	11	6.7	3.6	1.10
8	0.9	0.3	202	22	19	7.4	4.1	1.20
9	0.7	0.2	290	53	58	6.8	3.4	1.00
10	0.6	0.1	210	51	59	5.9	2.7	0.80
11	2.7	1.3	260	31	56	7.4	3.0	0.60
12	0.9	0.3	310	61	58	7.0	3.4	0.90
13	1.1	0.4	214	22	30	8.1	4.1	1.00
14	0.7	0.2	145	53	41	5.8	2.7	0.87
15	0.6	0.1	183	91	53	5.5	2.3	0.70
16	1.8	0.8	342	168	441	7.6	4.4	1.30
17	1.6	0.5	165	15	23	7.3	3.5	0.92
18	0.9	0.3	293	232	245	6.8	3.1	0.80
19	2.6	1.2	236	131	90	5.4	2.6	0.90
20	1.1	0.3	215	46	134	6.9	3.0	0.70
21	15.6	9.5	134	54	125	5.6	4.0	2.50
22	3.7	1.6	612	50	88	6.2	1.9	0.40
23	12.1	6.0	515	48	92	6.6	2.4	0.50
24	25.0	13.7	560	41	88	7.9	2.5	2.50
25	15.0	8.2	289	58	80	5.3	2.2	0.70

```
[0.81818181 0.72727275 0.80000001]
Liver Baseline: 78.18% (3.93%)
[1]
Negative
```

CANCER:

```
     Cell Thickness  Cell size  Cell shape  Marg.adesion  Epith.c.size  \
0                 5          1           1             1             2
1                 5          4           4             5             7
2                 3          1           1             1             2
3                 6          8           8             1             3
4                 4          1           1             3             2
..              ...        ...         ...           ...           ...
115              10         10          10             2            10
116               5          3           5             1             8
117               5          4           6             7             9
118               7          5           3             7             4
119               8          3           5             4             5

     Bare.nuclei  Bl.cromatin  Normal.nucleoli  Mitoses
0              1            3                1        1
1             10            3                2        1
2              2            3                1        1
3              4            3                7        1
4              1            3                1        1
..           ...          ...              ...      ...
115           10            5                3        3
116           10            5                3        1
117            7            8               10        1
118           10            7                5        5
119           10            1                6        2

[120 rows x 9 columns]
[0.91666669 0.91666669 1.         0.83333331 0.91666669 1.
 1.         1.         1.         0.83333331]
Cancer Baseline: 94.17% (6.51%)
[1]
Positive
```

CONCLUSION

We have introduced a simple and computationally efficient algorithm for gradient-based optimization of stochastic objective functions. Our method is aimed towards machine learningproblems with large datasets and/or high-dimensional parameter spaces. The method combinesthe advantages of two recently popular optimization methods: the ability of AdaGrad to deal with sparse gradients, and the ability of RMSProp to deal with non- stationary objectives. Themethod is straightforward to implement and requires little memory. The experiments confirm the analysis on the rate of convergence in convex problems. Overall, we found Adam to be robust and well-suited to a wide range of non-convex optimization problems in the field machine learning. The performance of Adam is acceptable and has promising results which nominate it for other optimization applications such as timetabling. Training the NN using theAdam for classifying many other datasets is an going research. Future work includes further analysis for the ADAM algorithm.

REFERENCES

1. IEEE International Conference on, pp. 6645–6649. IEEE, 2013. Hinton, G.E. and Salakhutdinov, R.R. Reducing the dimensionality of data with neural networks. Science, 313 (5786):504–507, 2006.
2. Duchi, John, Hazan, Elad, and Singer, Yoram. Adaptive subgradient methods for online learning and stochastic optimization. The Journal of Machine Learning Research, 12:2121–2159, 2011.
3. Moulines, Eric and Bach, Francis R. Non-asymptotic analysis of stochastic approximationalgorithms for machine learning. In Advances in Neural Information Processing Systems, pp. 451–459, 2011.
4. Krizhevsky, Alex, Sutskever, Ilya, and Hinton, Geoffrey E. Imagenet classification with deep convolutional neural networks. In Advances in neural information processing systems,pp. 1097–1105, 2012.
5. Deng, Li, Li, Jinyu, Huang, Jui-Ting, Yao, Kaisheng, Yu, Dong, Seide, Frank, Seltzer, Michael, Zweig, Geoff, He, Xiaodong, Williams, Jason, et al. Recent advances in deep learning for speech research at microsoft. ICASSP 2013, 2013.
6. Graves, Alex. Generating sequences with recurrent neural networks. arXiv preprint arXiv:1308.0850, 2013. Graves, Alex, Mohamed, Abdel-rahman, and Hinton, Geoffrey. Speech recognition with deep recurrent neural networks. In Acoustics, Speech and Signal Processing (ICASSP), 2013
7. Kingma, Diederik P and Welling, Max. Auto-Encoding Variational Bayes. In The 2ndInternational Conference on Learning Representations (ICLR), 2013.
8. Pascanu, Razvan and Bengio, Yoshua. Revisiting natural gradient for deep networks.arXiv preprint arXiv:1301.3584, 2013.
9. Tieleman, T. and Hinton, G. Lecture 6.5 - RMSProp, COURSERA: Neural Networks for Machine Learning. Technical report, 2012.
10. Zeiler, Matthew D. Adadelta: An adaptive learning rate method. arXiv preprint arXiv:1212.5701, 2012.
11. Kevin Bache and Moshe Lichman. 2013. UCI machine learning repository. (2013).