Predicting Autism Spectrum Disorder in Toddlers Using A Neural Network

Connor Sparling
Kevin Wong

Table of Contents

Introduction

- Autism is complex and varied.
- Symptoms may not be recognized right away
- Getting a diagnosis takes time
- The earlier the diagnosis, the better the outcome

Objective

- Train an artificial neural network to be able to classify Autism Spectrum Disorder with an emphasis on information from the Quantitative Checklist for Autism in Toddlers (Q-CHAT).

Data: Q-Chat Scores

- For questions 1 - 10:
  - If yes is answered for each question, add 1 to the total Q-Score.
  - Scores range from 0 - 10

Variable in Dataset	Corresponding Q-chat-10-Toddler Features
A1	Does your child look at you when you call his/her name?
A2	How easy is it for you to get eye contact with your child?
A3	Does your child point to indicate that s/he wants something? (e.g. a toy that is out of reach)
A4	Does your child point to share interest with you? (e.g. poin9ng at an interes9ng sight)
A5	Does your child pretend? (e.g. care for dolls, talk on a toy phone)
A6	Does your child follow where you're looking?
A7	If you or someone else in the family is visibly upset, does your child show signs of wan9ng to comfort them? (e.g. stroking hair, hugging them)
A8	Would you describe your child's first words as:
A9	Does your child use simple gestures? (e.g. wave goodbye)
A10	Does your child stare at nothing with no apparent purpose?

Data

- Database contains 1,054 instances.
- 18 Attributes including class variable.
- Questions A1 - A10
  - Binary [0, 1]
- Age:
  - Number of months old
- Sex:
  - String [‘Male’ or ‘Female’]
- Ethnicity:
  - String [‘White European’]
- Born with Jaundice:
  - Boolean [‘Yes’ or ‘No’]
- Family Member with ASD History
  - Boolean [‘Yes’ or ‘No’]
- Who is completing the test
  - String [‘Parent, ‘Self’, ‘Caregiver’, ‘Medical Staff’, ‘Clinician’, etc.]
- Class Variable:
  - String [‘Yes’ or ‘No’]

Case_No	A1	A2	A3	A4	A5	A6	A7	A8	A9	A10	Age_Mon	Qchat-10-	Sex	Ethnicity	Jaundice	Family_m	Who com	Class/ASD Traits
1	0	0	0	0	0	0	1	1	0	1	28	3	f	middle ea	yes	no	family me	No

Approach

- 2 Class approach [Yes, No] based on whether the person is in the spectrum or not.
- A FeedForward Backpropagation neural network was developed using Python and Keras.
  - 27 Input Nodes, 15 Hidden Nodes, and 1 Output Node
  - 75% Training Data, 25% Test Data
- A Multilayer Perceptron neural network was developed using Python and sklearn.
  - 17 Input Nodes, 50 Hidden Nodes, 2 Output Nodes
  - 75% Training Data, 25% Test Data
- Classifiers (KNN, Naive Bayes, Gradient Boosting, Random Forest) used to compare against model.

Results (feedforward network)

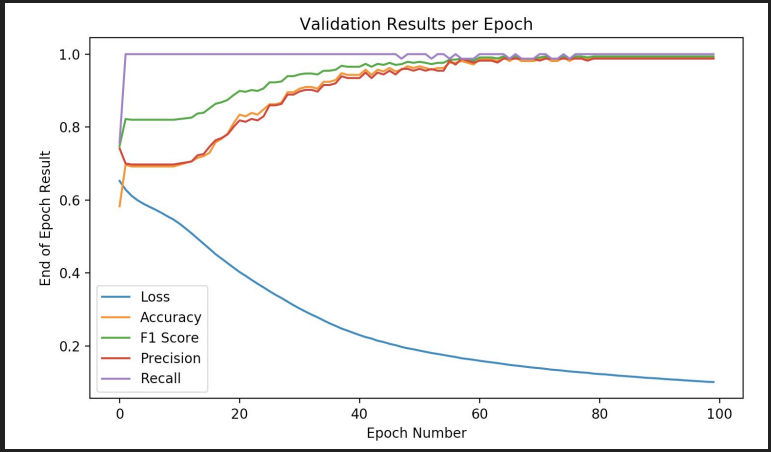
Loss: 0.052644977

Accuracy: 99.53%

F1 Score: 99.63%

Precision: 99.29%

Recall: 100.00%



Relevance of Data to Autism Classification:

1: QchatScore
2: A6
3: EthnicityIsNativeIndian
4: FamilyASD
5: EthnicityIsOthers
6: EthnicityIsLatino
7: Jaundice
8: A9
9: A8
10: A1
11: A10
12: A3
13: Age
14: A7
15: EthnicityIsPacifica
16: A5
17: SexIsM
18: EthnicityIsMixed
19: EthnicityIsAsian
20: EthnicityIsSouthAsian
21: EthnicityIsHispanic
22: EthnicityIsMiddleEastern
23: EthnicityIsBlack
24: SexIsF
25: EthnicityIsWhiteEuropean
26: A2
27: A4

Results (MLP)

Loss: 0.163690883

Accuracy: 99.52%

F1 Score: 99.68%

Precision: 99.84%

Recall: 99.36%

Confusion Matrix:

```
[[ 53   0]
 [  1 157]]
```

Classification Report:

	precision	recall	f1-score	support
0	0.98	1.00	0.99	53
1	1.00	0.99	1.00	158
accuracy			1.00	211
macro avg	0.99	1.00	0.99	211
weighted avg	1.00	1.00	1.00	211

Results (Classifiers)

KNN and NB

- Consistently have a few misclassified.
- Accuracy of high 90s

```
KNN-------------------------------------------------
Confusion Matrix:
[[ 66   5]
 [  6 134]]
Classification Report:
              precision    recall  f1-score   support

          No       0.92      0.93      0.92        71
         Yes       0.96      0.96      0.96       140

    accuracy                           0.95       211
   macro avg       0.94      0.94      0.94       211
weighted avg       0.95      0.95      0.95       211

Naive Bayes-----------------------------------------
Confusion Matrix:
[[ 67   4]
 [  1 139]]
Classification Report:
              precision    recall  f1-score   support

          No       0.99      0.94      0.96        71
         Yes       0.97      0.99      0.98       140

    accuracy                           0.98       211
   macro avg       0.98      0.97      0.97       211
weighted avg       0.98      0.98      0.98       211
```

Results (Classifiers and Feature Importance)

Random Forest----------

Confusion Matrix:

```
[[ 71   0]
 [  0 140]]
```

Classification Report:

	precision	recall	f1-score	support
No	1.00	1.00	1.00	71
Yes	1.00	1.00	1.00	140
accuracy			1.00	211
macro avg	1.00	1.00	1.00	211
weighted avg	1.00	1.00	1.00	211

Feature Importance:

qChat	0.547084
a9	0.089952
a7	0.069426
a6	0.063549
a5	0.051991
a1	0.051740
a2	0.038596
a8	0.026314
a4	0.020944
age	0.009911
a3	0.008650
ethnicity	0.007611
a10	0.006050
jaundice	0.002869
sex	0.002816
familyASD	0.001916
whoCompleted	0.000581

dtype: float64

Gradient Boosting----------

Confusion Matrix:

```
[[ 71   0]
 [  0 140]]
```

Classification Report:

	precision	recall	f1-score	support
No	1.00	1.00	1.00	71
Yes	1.00	1.00	1.00	140
accuracy			1.00	211
macro avg	1.00	1.00	1.00	211
weighted avg	1.00	1.00	1.00	211

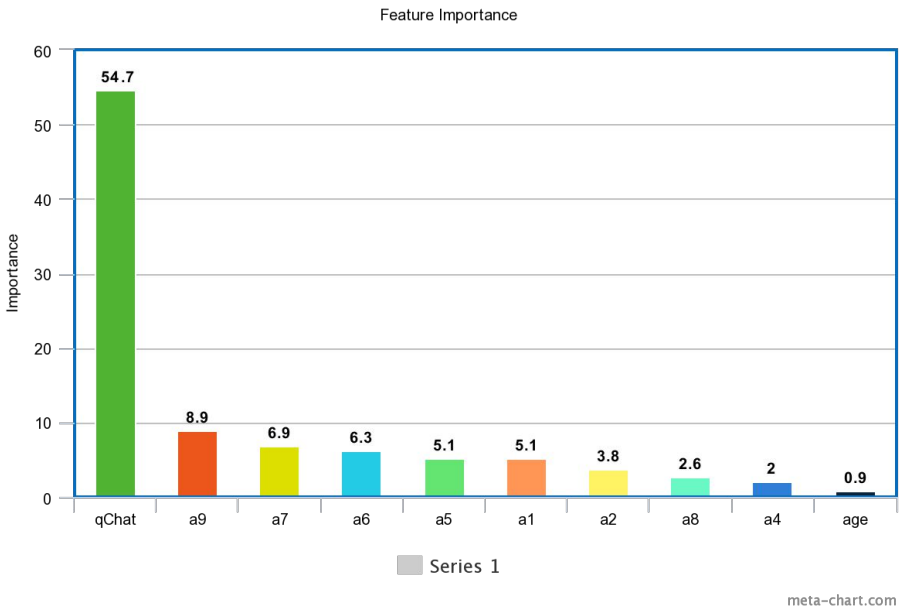
Feature Importance:

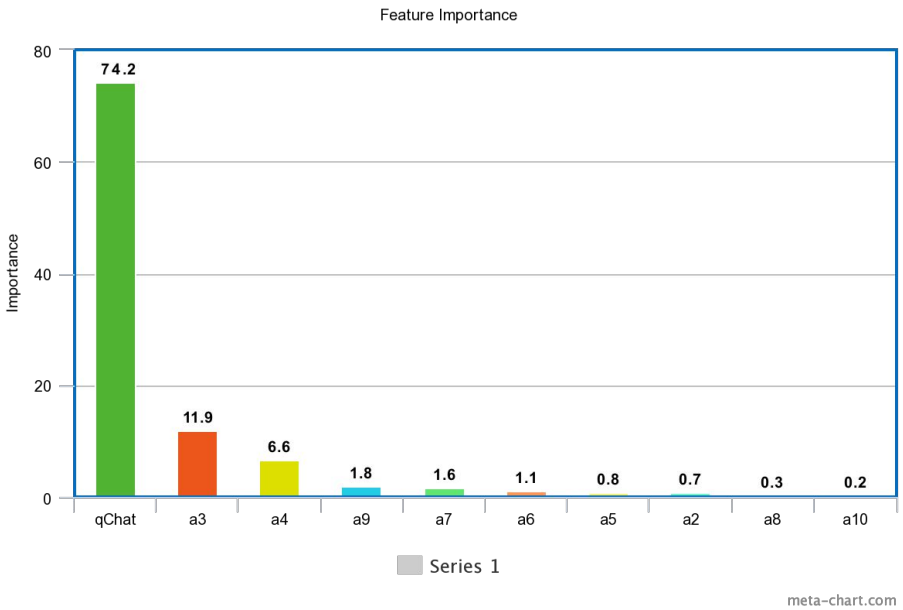
qChat	0.742287
a3	0.119315
a4	0.066467
a9	0.018462
a7	0.016596
a6	0.011690
a5	0.008831
a2	0.007156
a8	0.003050
a10	0.002236
a1	0.001793
ethnicity	0.001398
age	0.000576
jaundice	0.000103
sex	0.000020
whoCompleted	0.000018
familyASD	0.000000

dtype: float64

Feature Importance

- Q-Chat Scores are consistently at the top for feature importance.





Discussion

- Preliminary diagnosis of ASD can reduce the the amount of work a medical professional does and save time.
- The network can be used not just on toddlers but for adults as well.

Future Works

- Different severity levels for questions A1 - A10.
- Different severity levels for ASD
- More Questions