

Outcome predictors in autism spectrum disorders preschoolers undergoing treatment as usual: insights from an observational study using artificial neural networks

Antonio Narzisi¹
Filippo Muratori^{1,2}
Massimo Buscema^{3,4}
Sara Calderoni¹
Enzo Grossi^{3,5}

¹Department of Developmental Neuroscience, IRCCS Stella Maris Foundation, ²Department of Clinical and Experimental Medicine, University of Pisa, Pisa, Italy; ³Semeion Research Centre of Sciences of Communication, Rome, Italy; ⁴Department of Mathematical and Statistical Sciences, University of Colorado Denver, Denver, CO, USA; ⁵Autism Research Unit, Villa Santa Maria Institute, Tavernerio, Italy

Background: Treatment as usual (TAU) for autism spectrum disorders (ASDs) includes eclectic treatments usually available in the community and school inclusion with an individual support teacher. Artificial neural networks (ANNs) have never been used to study the effects of treatment in ASDs. The Auto Contractive Map (Auto-CM) is a kind of ANN able to discover trends and associations among variables creating a semantic connectivity map. The matrix of connections, visualized through a minimum spanning tree filter, takes into account nonlinear associations among variables and captures connection schemes among clusters. Our aim is to use Auto-CM to recognize variables to discriminate between responders versus no responders at TAU.

Methods: A total of 56 preschoolers with ASDs were recruited at different sites in Italy. They were evaluated at T0 and after 6 months of treatment (T1). The children were referred to community providers for usual treatments.

Results: At T1, the severity of autism measured through the Autism Diagnostic Observation Schedule decreased in 62% of involved children (Response), whereas it was the same or worse in 37% of the children (No Response). The application of the Semeion ANNs overcomes the 85% of global accuracy (Sine Net almost reaching 90%). Consequently, some of the tested algorithms were able to find a good correlation between some variables and TAU outcome. The semantic connectivity map obtained with the application of the Auto-CM system showed results that clearly indicated that “Response” cases can be visually separated from the “No Response” cases. It was possible to visualize a response area characterized by “Parents Involvement high”. The resultant No Response area strongly connected with “Parents Involvement low”.

Conclusion: The ANN model used in this study seems to be a promising tool for the identification of the variables involved in the positive response to TAU in autism.

Keywords: autism spectrum disorders, treatment, intervention, artificial neural networks, outcome

Introduction

Autism spectrum disorders (ASDs) encompass a broad spectrum of heterogeneous neurodevelopmental disorders characterized by social communication impairments and restricted repetitive patterns of behavior.¹

Recent studies have compared specific early manualized interventions versus treatment as usual (TAU) that is usually available in the communities.^{2–8} These studies were randomized controlled trials (RCTs), which are considered as the “gold standard” of the evidence-based research.⁹ Nevertheless, the debate on RCTs remains much discussed.

Correspondence: Antonio Narzisi
Department of Developmental Neuroscience, IRCCS Stella Maris Foundation, Via dei Giacinti 2, I-56018 Calambrone, Pisa, Italy
Tel +39 050 88 6308
Fax +39 050 88 6290
Email antonio.narzisi@fsm.unipi.it

One of the biggest problems associated with RCT studies is their distance from the real-world environment. The present dilemma raises a question whether we should use randomized trials or observational studies to assess the outcome of a particular disease such as autism. This question is really fundamental since an observational study might constitute the ideal medium for the application of artificial adaptive systems (AASs).

The central strength of an RCT is that groups of patients allocated to each treatment tend to be comparable. In addition, randomization leads to robust methods of hypothesis testing that requires a few statistical assumptions. For these reasons, RCT is often regarded as the “gold standard” of therapeutic and diagnostic research.

However, in real life, patients are not randomly assigned to receive manualized treatment given in a rigid, standardized way, as is the case in most RCTs.

Since, traditionally, the drawback of observational studies is the poor internal validity, in the recent years efforts have been made to develop improved methods to evaluate therapeutic effectiveness in the framework of observational studies.^{10–12}

AASs can analyze real-world data very efficiently. The internal validity of their assessment is provided by uniquely severe validation protocols, seldom used in classical statistics.^{13–15}

In the last 20 years, artificial neural networks (ANNs) have been used in the field of autism to investigate the mechanisms of developmental regression,¹⁶ to identify peculiar features in reach-and-throw movements,¹⁷ to predict the diagnosis,¹⁸ to study attention shift,¹⁹ and to discriminate children with autism from children with mental retardation.²⁰

We performed the present study to investigate whether this revolutionary mathematical approach can increase our knowledge on the connections among those variables in subjects who respond positively to TAU and hence identify the key variables to discriminate responders from nonresponders.

To accomplish this, we applied ANNs and other machine learning systems to assess their predictive capacity in distinguishing consistently the two outcomes of interest (Response vs No Response) of TAU and to identify the variables expressing the maximal amount of relevant information for this distinction.

ANNs allow a method of forecasting with an understanding of the relationship among variables, and in particular nonlinear relationships.^{11–22} ANNs function by initially learning a known set of data from a given problem with a known solution (training) and then the networks, inspired

by the analytical processes of the human brain, are able to reconstruct the imprecise rules, which may be underlying a complex set of data (testing).

Moreover, we used the Auto Contractive Map (Auto-CM), a special kind of ANN able to define the strength of the associations of each variable with all the others and to visually show the map of the main connections of the variables and the basic semantic of their ensemble.

Materials and methods

Population

In this work, we have explored in a new way some of the data from our previous study on brief outcome of children with autism under early treatment.²³ For this exploration, the sample consisted of 56 children (47 males, 9 females; mean age: 36.01 ± 0.79 months; age range: 18–60 months) with a DSM-IV-TR diagnosis of autistic disorder ($n=46$) or pervasive developmental disorder not otherwise specified ($n=10$). A total of 51 children received an Autism Diagnostic Observation Schedule (ADOS) classification of autism and 5 received an autism spectrum diagnosis. The mean non-verbal development quotient was 73.8 ± 18.3 (range: 50–125), and the mean general quotient was 59.1 ± 11.8 (range: 34–85). All the children were re-evaluated after 6 months and divided in two groups: responders and nonresponders.

Measurements

The assessment protocol was composed by gold standard measures: ADOS-Generic (the first author, AN, was certified to administer ADOS in clinical and research setting at the University of Michigan Autism Communication Centre; all the clinicians involved in this study were trained to administer ADOS in a clinical and research setting) and Griffiths Mental Developmental Scales and Vineland Adaptive Behavior Scales-II. We also used parent reports: MacArthur Communicative Development Inventories, Child Behavior Checklist (CBCL) 1½–5, and Parenting Stress Index (PSI). A detailed description of these assessment protocol was reported in our original study.²³

Procedure

The children were evaluated at T0 and after 6 months of treatment (T1). At T0, child clinical measures were well equable across treatment sites.

Intervention

All the children received TAU. It includes eclectic treatments usually available in the community and school inclusion

with an individual support teacher. TAU included speech therapy and/or psycho-educative therapy. Each child's program comprises individual objectives but is mainly based on therapist expertise rather than on manualized treatment protocols or uniform training. Treatments can be placed within a continuum ranging from highly structured behavioral approaches to approaches that follow the interests of the child in a naturalistic setting and are based on a developmental curriculum in a relational-based context (a deep explanation of TAU is also reported in our original paper).²³

Outcome

The primary outcome was the ADOS calibrated severity score (ADOS-CSS) in order to distinguish children who positively respond to treatment (hereinafter Response) versus nonresponders (hereinafter No Response). ADOS-CSS is a measurement of the severity of the autism symptoms. The ADOS-CSS scores had more uniform distributions across developmental groups and were less influenced by participant demographics than raw totals. This metric is useful in comparing assessments across time and identifying trajectories of autism severity for clinical research.²⁴

Mathematical methods

To evaluate the possibility to predict the treatment outcome (Response vs No Response) using as input data all the 25 variables on study (Table 1) we have trained different machine learning systems available on WEKA data mining software (University of Waikato, Hamilton, New Zealand)^{25–27} and on Semeion Research Centre depository, Rome, Italy, as

classification tools to predict the treatment outcome using the Training and Testing validation protocol. This protocol has been described in detail elsewhere.^{14,15}

The learning machines algorithms developed at the University of Waikato, New Zealand, available on the WEKA data mining software are listed in Table 2,^{28–34} whereas two ANNs (Self Momentum Back Propagation and Sine Net)^{35,36} were implemented in "Supervised ANNs Software", developed at the Semeion Research Center (Buscema M; Supervised ANNs. Semeion software #12, version 16.0).

However, since noisy input attributes sometimes can hide the small meaningful information embedded in other attributes, a pruning procedure was used as a preprocessing tool to eliminate noisy variables before the outcome prediction of the main test. In order to conduct that procedure, a special and powerful recently published input selection algorithm named Training With Input Selection and Testing (TWIST) was applied^{37–44} and developed in a special research software at the Semeion Research Center (Buscema M [2006–2012] TWIST Input Search, Semeion software #39, version 3.2).

TWIST algorithm

As described in the work by Coppédè,²¹ the TWIST algorithm is a complex algorithm that is able to search for the best distribution of the global dataset divided in two optimally balanced subsets containing a minimum number of input features useful for optimal pattern recognition. TWIST is an evolutionary algorithm based on a seminal paper about genetic doping systems, already applied to medical data with very promising results.^{11,22,26,38–44} TWIST selected 9 of the original attributes (Table 3) and generated a global dataset of 25 attributes, and 2 optimal subsets for training and testing. We then applied the K-Fold protocol to the global dataset to verify whether the nine attributes selected by TWIST may improve the performances of the learning machines already applied to the original dataset. Moreover, as a second step, we applied the same learning machines to the two subsets generated directly by TWIST.

Table 1 Variables on study

Age	PSI Total
Sex	PSI Child Domain
Griffiths locomotor	PSI Parent Domain
Griffiths personal	Vineland Communication
Griffiths speech	Vineland Daily Living
Griffiths eye	Vineland Socialization
Griffiths performance	Vineland Motor
Griffiths general	Vineland Composite
CBCL-int	Hours of Treatment
CBCL-ext	Parents Involvement
CBCL-tot	Treatment Centre Localization
MacArthur Comprehension	Response (ADOS-CSS)
MacArthur Expressive	No Response (ADOS-CSS)
MacArthur Gestures	

Notes: Griffiths locomotor is locomotor development; Griffiths personal is personal-social development; Griffiths speech is hearing and speech; Griffiths eye is hand and eye coordination; Griffiths general is general quotient.

Abbreviations: CBCL, Child Behavior Checklist; int, internalizing; ext, externalizing; tot, total; PSI, Parenting Stress Index; ADOS-CSS, Autism Diagnostic Observation Schedule-calibrated severity score.

Table 2 Learning machines in the WEKA software package

Learning machine	Nick name	References
Logistic	Logistic	Hosmer, Lemeshow S ²⁸
J48	LogitBoost	Ross ²⁹
Multilayer perceptron	MLP	Collobert, Bengio ³⁰
Naïve Bayes	NaivBayes	George, Langley ³¹
Random forest	RandomForest	Livingston ³²
Rotation forest	RotationForest	Rodriguez et al ³³
Sequential minimal optimization	SMO	Keerthi, Gilbert ³⁴

Table 3 Variables selected by the TWIST system

CBCL-int
CBCL-ext
CBCL-tot
PSI Child Domain
PSI Parent Domain
Vineland Communication
Vineland Motor
Vineland Composite
Parents Involvement

Abbreviations: TWIST, Training With Input Selection and Testing; CBCL, Child Behavior Checklist; int, internalizing; ext, externalizing; tot, total; PSI, Parenting Stress Index.

Semantic connectivity map

An existing mapping method^{45,46} was used to highlight through a graph the most important links among variables, using a mathematical approach called Auto-CM. Auto-CM is a special kind of ANN able to find the consistent patterns and/or systematic relationships among variables.^{45,46} Auto-CM ANN was designed by Buscema M at the Semeion Research Center, and developed in specific research softwares (AutoCM – Auto Contractive Map, Semeion software #46, version 6.0; Modular Auto-Associative ANN, Semeion software #51, version 18.1).

Auto-CM can also recognize in hard conditions, that is, when the connections of the main diagonal of the second connections matrix are removed. When the learning process is organized in this way, Auto-CM seems to find specific relationships between each variable and any other. Consequently, from an experimental point of view, it seems that the ranking of its connections matrix is equal to the ranking of the joint probability between each variable and the others. For the Auto-CM analysis, the same 25 variables used for predictive analysis were employed, except for sex and treatment center localization. We transformed the 23 input variables in 46 input variables constructing for each of the variable, scaled from 0 to 1, its complement as explained in a previous paper.⁴⁷

In the complement transformation, by subtracting the scaled value from 1, the system was allowed to project and point out the fuzzy position of each variable according to its low values. This is important because in nonlinear systems, the position of high and low values of a given variable is not necessarily symmetric.

In this way, the projection of the original variables tended to show high values, whereas the complement transformation tended to show low values of the original variables. In the map, we have named these two different forms as high and low. This preprocessing scaling is necessary to make possible a proportional comparison among all the variables

and to understand the existing links of each variable when the values tend to be high or low.

Results

Response vs No Response

At T1, ADOS-CSS improved in 35 (62.5%) of the 56 children (Response), whereas it was the same or worse in 21 (37.5%) of the 56 children (No Response).

In Table 4, the independent *t*-test and Cohen's *d* effect size results of the comparison between Response and No Response groups at T0 assessment are shown. There were significant differences at CBCL (Internalizing Problems) and at PSI (Total and Child Domains).

Prediction of the outcome with machine learning algorithms

Tables 5 and 6 show the results in the two selected strategies of prediction (with and without variable selection, respectively).

Using all the 25 variables in the dataset as input vectors, the classification capabilities of all the algorithms are rather low, except the Sine Net and Back Propagation (77.35% and 77.99% of global accuracy, respectively). The conclusion from Table 5 could be that there is a moderate evidence of correlation between these variables and TAU outcome. However, the application of the TWIST algorithm to eliminate noisy variables before the main test of pattern recognition allowed the selection of nine attributes (listed in Table 3). Most of the learning machines improve their performances dramatically (up to 80% and more of global accuracy), and both the Semeion ANNs overcome 85% of global accuracy (Sine Net almost reaching 90%) (Table 6). Consequently, some of the tested algorithms found a good correlation between some variables and TAU outcome, once the noisy attributes were removed (see Supplementary materials for explanation of different machine learning).

Semantic connectivity map

Figure 1 reports the semantic connectivity map. As described by Coppèdè,²¹ in order to better understand the meaning of the connections, a numerical value is applied to each edge of the graph. This value, deriving from the original weight developed by Auto-CM during the training phase scaled from 0 to 1, is proportional to the strength of the connections between two variables. Moreover, by means of Auto-CM, it is possible to obtain not only the direction of the association as provided by standard statistical analyses but also specifically the strength of this association (link strength [LS]).

Table 4 Comparison at T0 among Response vs No Response groups

Demographic and clinical characteristics	Response (T0)		No response (T0)		Independent t-test			Effect size
	Mean	SD	Mean	SD	t	df	P-value	Cohen's d
Age								
Years	2.90	0.84	3.17	0.66	0.29	54	NS	0.07
ADOS								
Calibrated severity score	7.54	2.10	7.23	1.44	0.58	54	NS	0.15
Griffiths								
GQ	59.08	11.05	59.33	13.41	0.07	54	NS	0.01
Locomotor development	83.09	16.5	117.28	20.2	2.27	54	NS	0.61
Personal-social development	49.65	19.13	56.19	15.19	1.33	54	NS	0.36
Hearing and speech	33.91	18.63	36.57	13.31	0.57	54	NS	0.15
Hand and eye coordination	60.62	13.62	62.14	17.35	0.36	54	NS	0.09
Performance test	73.42	17.03	74.42	20.85	0.19	54	NS	0.05
Vineland-II								
Composite	68.58	13.23	62.05	9.35	1.98	54	NS	0.53
Communication	63.22	13.41	52.90	15.99	2.59	54	NS	0.70
Daily Living	72	14.65	66.95	11.78	1.33	54	NS	0.36
Socialization	63.37	13.27	64.14	8.08	0.24	54	NS	0.06
Motor	82.37	16.97	76.57	10.02	1.42	54	NS	0.38
MacArthur								
Production	52.45	75.56	58.71	92.39	0.27	54	NS	0.07
Comprehension	130.6	97.99	137.14	101.55	0.23	54	NS	0.06
Gestures	27.8	12.06	28.7	16.11	0.25	54	NS	0.06
CBCL								
Internalizing problems	68.34	12.19	56.80	8.61	3.79	54	P<0.001	1.03
Externalizing problems	61.6	13.50	51.95	9.46	2.87	54	NS	0.78
Total problems	67.4	15.03	55.80	15.13	3.12	54	NS	0.84
PSI								
Total	304.71	67.76	244.23	38.35	3.73	54	P<0.001	1.01
Child Domain	146.82	30.23	119.80	20.01	3.63	54	P<0.001	0.98
Parent Domain	157.6	41.70	125.66	21.93	3.24	54	NS	0.88

Abbreviations: SD, standard deviation; ADOS, Autism Diagnostic Observation Schedule; GQ, General Quotient; CBCL, Child Behavior Checklist; PSI, Parenting Stress Index; NS, nothing significant.

It was possible to visualize a Response area characterized by “Parents Involvement high” (LS=0.98) and “MacArthur Expressive low” (LS=0.99).

This last condition was linked to: “Age low” (LS=0.99), “Vineland Composite low” (LS=0.99), “MacArthur Comprehension low” (LS=0.99), and “Griffith Locomotor low” (LS=1.00). Globally, all Griffiths scales, linked to

“Response”, showed low scores: Personal, Speech, Eye, Performance, and General.

Otherwise, the resultant No Response area was highly connected only with “Parents Involvement low” (LS=0.98). This condition was directly linked to “PSI total low” (LS=0.99), which was linked to low scores on CBCL scales.

Table 5 Predictive results without variable selection

Machine learning systems	Response	No response	Global accuracy	Errors
FF_Bp	74.17%	81.82%	77.99%	12
FF_Sn	68.33%	86.36%	77.35%	13
NaivBayes	70.00%	72.73%	71.36%	16
RotationForest	88.33%	51.82%	70.08%	14
SMO	76.67%	61.82%	69.24%	16
RandomForest	85.83%	38.64%	62.23%	18
J48	75.83%	47.73%	61.78%	19
MLP	72.50%	46.82%	59.66%	20
Logistic regression	70.00%	47.27%	58.64%	21

Note: The results are the average of two testing experiments with training-testing A-B and B-A sequences.

Abbreviations: FF_Bp, feed forward Back Propagation; FF_Sn, feed forward Sine Net; MLP, multilayer perceptron; SMO, sequential minimal optimization.

Table 6 Predictive results with variable selection

Machine learning systems	Response	No response	Global accuracy	Errors
FF_Sn	93.18%	85.29%	89.24%	8
FF_Bp	86.36%	85.29%	85.83%	11
IBk	88.57%	80.95%	84.76%	8
NaivBayes (WEKA)	68.57%	76.19%	72.38%	16
SMO	77.14%	66.67%	71.90%	15
NaivBayes	57.87%	85.29%	71.58%	18
MLP	80.00%	57.14%	68.57%	16
Logistic regression	77.14%	52.38%	64.76%	18

Note: The results are the average of two testing experiments with training-testing A-B and B-A sequences.

Abbreviations: FF_Sn, feed forward Sine Net; FF_Bp, feed forward Back Propagation; IBk, instance-based learning algorithm; MLP, multilayer perceptron; SMO, sequential minimal optimization.

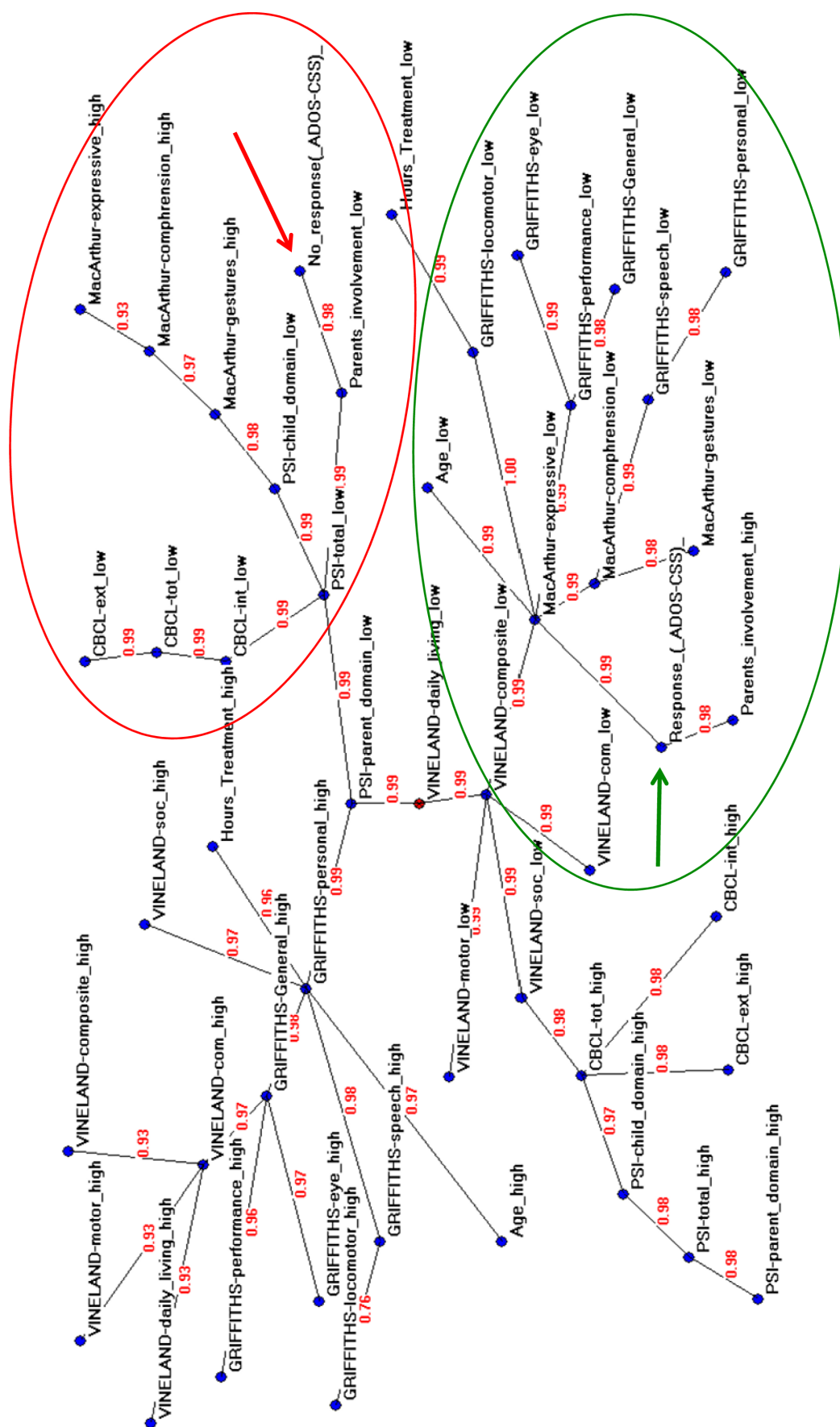


Figure 1 Semantic connectivity map obtained with Auto-Cm System.

Notes: The figures on the arcs of the graph refer to the strength of the association between two adjacent nodes. The range of this value is from 0 to 1. Red arrow points to the no response group; green arrow points to the response group.

Abbreviations: ADOS-CSS, Autism Diagnostic Observation Schedule-Calibrated Severity Score; CBCL, Child Behavior Checklist; int, internalizing; ext, externalizing; tot, total; Griffiths (locomotor, Locomotor development; personal, Personal-social development; speech, Hearing and eye coordination; general, General quotient); PSI, Parenting Stress Index; Vineland (Com, Communication; Daily Living, Daily Living Skills; Soc, Socialization).

In general, “No Response” area was linked to low PSI scores: both on Parent Domain and Child Domain, and high MacArthur scores (Expressive, Comprehension, and Gestures).

Discussion

The present study represents the first attempt to use ANNs in the arena of the research on ASD treatment. Our aim was to see whether ANNs were able to discriminate children who responded positively to TAU in terms of reduction of autism severity, using a set of variables describing behavioral, developmental and adaptive level profiles, and parental distress.

Despite the observational nature of the study, thanks to ANNs capacity, it was possible to build a predictive model of outcome response, an objective which could not be reached in our previous research work.²³ In fact, through the TWIST system, we established a consistent possibility to predict the status of being a responder or a nonresponder on the basis of nine variables (selected out of 25), which allowed to reach up to 89% global accuracy to some of the used learning machines. These selected variables contain specific information to discriminate between the two responder conditions. It was unexpected that, among these predictors, cognitive and language levels were not present. Most studies in fact have indicated that children with lower IQ are less likely to undergo positive gains.^{48,49} However, other studies have clearly demonstrated that, even among children with equally impaired cognition and language, individual response to the same treatment often differ markedly.⁵⁰ According to this latter finding, this study suggested that other factors not unique to ASD, such as parent involvement and stress, may be better predictors of treatment outcomes.

The semantic connectivity map obtained by means of the Auto-CM system has identified parent involvement as the main variable that influences the positive outcome of children under treatment; on the other side, no parent involvement is the main factor predicting negative outcomes. This finding, although partially expected,^{50–57} underlines the importance of involving parents who no longer have to be “left out” of the treatment room. Interestingly, a recent comprehensive synthesis of existing meta-analyses of Early Intensive Behavioral Intervention for young children with ASD published from 2009 to 2011 reported parent inclusion as a crucial factor for enhancing treatment effectiveness.⁵⁵

First, parents must be viewed as important participants in the intervention, and therapist-delivered treatment programs must be accompanied by parent-training methods.⁵⁶ In fact, this tenant has continued as part of the most recent approaches to early intervention in autism.⁵⁷ Second, this result is on the same

wavelength with findings of a recent meta-analysis that support the positive impact of psychosocial interventions delivered by nonspecialist providers as well as the parents of children with ASD.⁵⁸ Finally, the positive effect of parent involvement during therapy makes it necessary in the future to assess parent–child interaction as a possible outcome measure.⁵⁹

In addition to the direct involvement of parents, semantic connectivity map has identified other predictors of better outcome in terms of reduction in the severity of autism after TAU.

First, the young age in which the child begins treatment is consistent with the finding that confirms others research works that have underlined the importance of young age at the start of the treatment as a factor to promote benefits in the social communication domain.^{60–65} According to these authors, it is largely hypothesized that the better outcome might be due to the higher brain plasticity at this early age.⁶⁶

Second, young children are more likely to undergo positive gains if, at the beginning, they have low language and cognitive performances. Rogers⁶⁷ has already suggested, some years ago, that the evidence of direct links between pretreatment language abilities and treatment outcomes is contradictory. For example, Fenske⁶⁰ mentioned that the presence of language abilities not always predict positive outcomes in young treated children. The reason for this counterintuitive finding needs more investigations. It could be hypothesized that at this young age, a later development of language means that it is less interfered by the autistic process. It is possible that if language already has autistic features, other gains in the social/pragmatic language become more difficult. These children could be most resistant to change than children having low language performances when they started the treatment. On the contrary, if language develops during a sustained social-communicative program, it has more chances to have typical features and it could have cascading effects on global development.

Semantic connectivity map shows that cognitive functioning cannot be considered a critical factor affecting outcomes in young children with ASD.⁶⁸ Although some studies showed that having higher IQ at intake is predictive of a better social performances after treatment,⁶⁵ other studies found no relation between pretreatment IQ and outcomes.^{62,69} Thus, the role of the initial IQ as a predictor of outcome needs to be more investigated in future studies.

Third, the total number of hours of treatment was not predictive of better outcome. The intensiveness of treatment is a longstanding conflicting discussion point in the arena of autism treatment. Although some studies have described

best outcome when maximum hours per week of treatment is provided,⁷⁰ other studies, which specifically examined outcome effects of hours per week of treatment, have found no differences in benefits obtained.⁷¹ In any case, this study suggests that the concept of intensiveness should be reformulated taking into account which type of support children have outside specific hours of treatment. For example, parent involvement means that some part of treatment is provided by parents during everyday life, thereby increasing the hours of treatment.

Again the stronger variables influencing no response to treatment, in addition to low parental involvement during the treatment, are the low stress levels of parents and the low behavioral problems of the child.

Usually, a child with a diagnosis of autism could be a source of stress for the family⁷² and the parental stress could reach higher levels when the child begins the treatment.²⁴ On the contrary, the low level of parental stress could be linked to a low awareness of the severity of diagnosis of their children, so that these parents could be less active in being involved, seeking, and planning the treatment solutions for their children. The low stress could also be linked to the low level of child's behavioral problems that often represent one of the most significant sources of stress for the families.^{73–75} It is worth noting that a recent study⁷⁶ has reported that behavioral problems that are not core symptoms of ASD were associated with an high parental stress.

The low behavioral problems could indicate that a certain type of children are less sensitive to TAU: first of all, this behavioral pattern seems to describe the aloof type of autism spectrum, according to Wing,⁷⁷ that is, subjects with a total disengagement from social interaction and a failure to engage in interpersonal reciprocity; second, these patients seem to be free of regulation disorder and/or anxious or opposite comorbidity frequently reported in ASD.^{78,79} Our hypothesis is that the absence of these comorbid features could mean a more rigid and less treatable autism. These children could be most resistant to change than children having dysregulatory comorbid pattern or simply they are less sensitive to TAU and need a different type of treatment.

Strengths and limitations

The observational approach combined with the use of ANNs represents the main point of strength of this study. Cases spontaneously arrived at clinics represent a real autistic population of preschoolers, which received treatments by their communities. This is a big advantage with respect to translational need of current clinical research. In this scenario,

although the lack of an RCT trial could be considered a weakness from a methodological point of view, the use of ANNs allowed us to overcome the main problem of observational design approach (ie, the low internal validity).

Special protocols of external validation methods, including cross-validation, and the dataset splitting into training and testing samples are able to increase the internal validity of clinical studies such as ours. Originally developed for neural network approaches, these validation protocols are now frequently applied to these traditional analyses. In this way, the use of ANNs is a powerful booster for the more widespread use of observational design in clinical research.

Moreover, ANN could be considered a more “naturalistic” approach than RCT in the field of autism research. In fact, in real life, patients are not randomly assigned to receive manualized treatment given in a standardized way, as is the case in most RCTs.

Patients with autism in the real world have comorbid conditions (ie, epilepsy; severe mental retardation) that normally would preclude them from entering an RCT, or they tend to be less compliant to the treatment and less subject to artificial expectations of recovery, arising from enthusiastic feedback from highly motivated investigators (Hawthorn effect).

An RCT tries to maintain a specific variable (the type of intervention) under control, thanks to randomization, presuming that all independent variables will be automatically balanced between treatment groups, and, therefore, the eventual differences on the outcome might be attributed to the treatment type. Unfortunately, the balance of independent variables at the group level may not be the same at the single individual level nor it allows for the discovery of an eventual complex interaction between independent and dependent variables.

Since translational research has to do with real life, one would be more interesting in “effectiveness” rather than “efficacy”.

Effectiveness tends to answer to the question that whether the intervention works in the real world. Although effectiveness is much more difficult to assess than efficacy, it is now recognized as being the most important factor in deciding whether a particular agent is worth the resources that it consumes.

Since traditionally the drawback of observational studies is the poor internal validity, in the recent years efforts have been made to develop improved methods to evaluate therapeutic effectiveness in the framework of observational studies.

AASs can analyze real-world data very efficiently and it is very important for the autism community. The internal

validity of their assessment is provided by uniquely severe validation protocols, seldom used in classical statistics.

The main limitation of this study is the relatively small sample size. The clinical applicability of ANNs should be tested in large, multicenter, prospective clinical trials on treatment effectiveness.

Moreover, although this study found some interesting predicting factors, it has not included many other potential predicting factors (eg, the features of the parents and the family, some biomarkers of the disease). To include all these, possible variables will be very important for a good prediction model. Hence, the current study is preliminary, as a methodological exploration on the path to accurate prediction.

In conclusion, the ANN model used in this study appears to be a promising tool for the identification of the variables involved in the positive or negative response to TAU in autism. The identification of these variables represents a core step to respond to the key question “what works for whom” and thus to pave the way for treatment personalization.

Acknowledgments

The study was funded by the Italian Ministry of Health (IDIA project, Inquiry into Disruption of Intersubjective Equipment in Autism Spectrum Disorder in Childhood). SC was partly supported by the Italian Ministry of Health and by Tuscany Region with the grant (GR-2010-2317873). The funders had no role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Disclosure

The authors report no conflicts of interest in this work.

References

1. American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders*. 5th ed. Arlington (VA): American Psychiatric Publishing; 2013.
2. Volkmar F, Siegel M, Woodbury-Smith M, et al. Practice parameter for the assessment and treatment of children and adolescents with autism spectrum disorder. *J Am Acad Child Adolesc Psychiatry*. 2014;53(2):237–257.
3. Kasari C. Are we there yet? The state of early prediction and intervention in autism spectrum disorder. *J Am Acad Child Adolesc Psychiatry*. 2013;53(2):133–134.
4. Dawson G, Bernier R. A quarter century of progress on the early detection and treatment of autism spectrum disorder. *Dev Psychopathol*. 2013; 25(4 Pt 2):1455–1472.
5. Ospina MB, Krebs Seida J, Clark B, et al. Behavioural and developmental interventions for autism spectrum disorder: a clinical systematic review. *PLoS One*. 2008;3(11):e3755. Epub 2008 Nov 18.
6. Narzisi A, Colombi C, Balottin U, Muratori F. Non-pharmacological treatments in autism spectrum disorders: an overview on early interventions for pre-schoolers. *Curr Clin Pharmacol*. Epub 2013 Sep 20.
7. Dawson G, Rogers S, Munson J, et al. Randomized, controlled trial of an intervention for toddlers with autism: the Early Start Denver Model. *Pediatrics*. 2012;125(1):e17–e23.
8. Green J, Charman T, McConachie H, et al. Parent-mediated communication-focused treatment in children with autism (PACT): a randomised controlled trial. *Lancet*. 2010;375(9732):2152–2160. Epub 2010 May 20.
9. Grossi E. Technology Transfer from the science of medicine to the real world: the potential role played by artificial adaptive systems. *Subst Use Misuse*. 2007;42(2):267–304.
10. Rutten-van Mölken MP, van Doorslaer EK, van Vliet RC. Statistical analysis of cost outcomes in a randomized controlled clinical trial. *Health Economics*. 1994;3:333–345.
11. Grossi E, Mancini A, Buscema M. International experience on the use of artificial neural networks in gastroenterology. *Dig Liver Dis*. 2007;39: 278–285.
12. Horwitz RI, Viscoli CM, Clemens JD, Sadock RT. Developing improved observational methods for evaluating therapeutic effectiveness. *Am J Med*. 1990;89(5):630–638.
13. Vomweg TW, Buscema M, Kauczor HU, et al. Improved artificial neural networks in prediction of malignancy of lesions in contrast-enhanced MR-mammography. *Med Phys*. 2003;30(9):2350–2359.
14. Andriulli A, Grossi E, Buscema M, et al. Contribution of artificial neural networks to the classification and treatment of patients with uninvestigated dyspepsia. *Dig Liver Dis*. 2003;35:222–231.
15. Mecocci P, Grossi E, Buscema M, et al. Use of artificial networks in clinical trials: a pilot study to predict responsiveness to Donepezil in Alzheimer's disease. *J Am Geriatr Soc*. 2002;50(11):1857–1860.
16. Thomas MS, Knowland VC, Karmiloff-Smith A. Mechanisms of developmental regression in autism and the broader phenotype: a neural network modeling approach. *Psychol Rev*. 2011;118(4):637–654.
17. Perego P, Forti S, Crippa A, Valli A, Reni G. Reach and throw movement analysis with support vector machines in early diagnosis of autism. *Conf Proc IEEE Eng Med Biol Soc*. 2009;2555–2558.
18. Arthi K, Tamilarasi A. Prediction of autistic disorder using neuro fuzzy system by applying ANN technique. *Int J Dev Neurosci*. 2008; 26(7):699–704. Epub 2008 Jul 26.
19. Gustafsson L, Papliński AP. Self-organization of an artificial neural network subjected to attention shift impairments and familiarity preference, characteristics studied in autism. *J Autism Dev Disord*. 2004;34(2):189–198.
20. Cohen IL. An artificial neural network analogue of learning in autism. *Biol Psychiatry*. 1994;36(1):5–20.
21. Coppèdè F, Grossi E, Migheli F, Migliore L. Polymorphisms in folate-metabolizing genes, chromosome damage, and risk of Down syndrome in Italian women: identification of key factors using artificial neural networks. *BMC Med Genomics*. 2010;3:42.
22. Penco S, Grossi E, Cheng S, et al. Assessment of the role of genetic polymorphism in venous thrombosis through artificial neural networks. *Ann Hum Genet*. 2005;69:693–706.
23. Muratori F, Narzisi A, IDIA Consortium. Exploratory study describing 6-months outcomes for young children with autism who receive treatment as usual (TAU) in Italy. *Neuropsychiatr Dis Treat*. 2014;8(10):577–586.
24. Gotham K, Pickles A, Lord C. Standardizing ADOS scores for a measure of severity in autism spectrum disorders. *J Autism Dev Disord*. 2009; 39(5):693–705. Epub 2008 Dec 12.
25. Hall M, Frank E, Holmes G, et al. The WEKA data mining software: an update. *SIGKDD Explorations*. 2009;11(1):10–18.
26. Buscema M, Grossi E, Intraligi M, et al. An Optimized experimental protocol based on neuro-evolutionary algorithms. Application to the classification of dyspeptic patients and to the prediction of the effectiveness of their treatment. *Artif Intell Med*. 2005;34:279–305.
27. Buscema M. Genetic Doping Algorithm (GenD): theory and application. *Expert Syst*. 2004;21:63–79.
28. Hosmer DW, Lemeshow S. *Applied Logistic Regression*. 2nd ed. Wiley; 2000.
29. Ross Q. *C4.5: Programs for Machine Learning*. San Mateo (CA): Morgan Kaufmann Publishers; 1993.
30. Collobert R, Bengio S. Links between perceptrons, MLPs and SVMs. Proc Int'l Conf on Machine Learning (ICML). ACM Digital Library 2004. Available at <http://dl.acm.org/citation.cfm?id=1015415>. Accessed 2 June 2015.

31. George HJ, Langley P. Estimating continuous distributions in Bayesian classifiers. In: Eleventh Conference on Uncertainty in Artificial Intelligence; 1995; San Mateo, CA, 338–345.
32. Livingston F. Implementing Breiman's random forest algorithm into Weka. ECE591Q Machine Learning Conference Papers; November 27; 2005.
33. Rodriguez JJ, Kuncheva LI, Alonso CJ. Rotation forest: a new classifier ensemble method. *IEEE Trans Pattern Anal Mach Intell*. 2006; 28(10):1619–1630.
34. Keerthi SS, Gilbert EG. Convergence of a generalized SMO algorithm for SVM. *Machine Learning*. 2002;351–360.
35. Buscema M. Back propagation neural networks. *Subst Use Misuse*. 1998;33:233–270.
36. Buscema M, Terzi S, Breda M. Using sinusoidal modulated weights improve feed-forward neural networks performances in classification and functional approximation problems. *WSEAS Trans Inf Sci Appl*. 2006;3:885–893.
37. Buscema M, Breda M, Lodwick W. Training With Input Selection and Testing (TWIST) algorithm: a significant advance in pattern recognition performance of machine learning. *J Intell Learn Syst Appl*. 2013;5:29–38.
38. Dietterich TG. Approximate statistical tests for comparing supervised classification learning algorithms. *Neural Comput*. 1998;10(7):1885–1924.
39. Lahner E, Intraligi M, Buscema M, et al. Artificial neural networks in the recognition of the presence of thyroid disease in patients with atrophic body gastritis. *World J Gastroenterol*. 2008;14:563–568.
40. Buri L, Hassan C, Bersani G, et al. Appropriateness guidelines and predictive rules to select patients for upper endoscopy: a nationwide multicenter study. *Am J Gastroenterol*. 2010;105:1327–1337.
41. Street ME, Grossi E, Volta C, Faleschini E, Bernasconi S. Placental determinants of fetal growth: identification of key factors in the insulin-like growth factor and cytokine systems using artificial neural networks. *BMC Pediatr*. 2008;8:24.
42. Buscema M, Grossi E, Capriotti M, Babiloni C, Rossini PM. The I.F.A.S.T. model allows the prediction of conversion to Alzheimer disease in patients with mild cognitive impairment with high degree of accuracy. *Curr Alzheimer Res*. 2010;7:173–187.
43. Rotondano G, Cipolletta L, Grossi E, et al. Artificial neural networks accurately predict mortality in patients with non variceal upper GI bleeding. *Gastrointest Endoscop*. 2011;73:218–226.
44. Pace F, Riegler G, de Leone A, et al. Is it possible to clinically differentiate erosive from non erosive reflux disease patients? A study using an artificial neural networks-assisted algorithm. *Eur J Gastroenterol Hepatol*. 2010;22:1163–1168.
45. Buscema M, Grossi E. The semantic connectivity map: an adapting self-organising knowledge discovery method in data bases. Experience in gastro-oesophageal reflux disease. *Int J Data Min Bioinform*. 2008; 2:362–404.
46. Buscema M, Grossi E, Snowdon D, Antuono P. Auto-Contractive Maps: an artificial adaptive system for data mining. An application to Alzheimer disease. *Curr Alzheimer Res*. 2008;5:481–498.
47. Gironi M, Saresella M, Rovaris M, et al. A novel data mining system points out hidden relationships between immunological markers in multiple sclerosis. *Immun Ageing*. 2013;10:1.
48. Howlin P, Moss P, Savage S, Rutter M. Social outcomes in mid- to later adulthood among individuals diagnosed with autism and average nonverbal IQ as children. *J Am Acad Child Adolesc Psychiatry*. 2013; 52(6):572–581. Epub 2013 Apr 24.
49. Trembath D, Balandin S, Togher L, Stancliffe RJ. Peer-mediated teaching and augmentative and alternative communication for preschool-aged children with autism. *J Intellect Dev Disabil*. 2009;34(2):173–186.
50. Vismara LA, Colombi C, Rogers SJ. Can one hour per week of therapy lead to lasting changes in young children with autism? *Autism*. 2009;13(1):93–115.
51. Anderson SR, Romanczyk RG. Early intervention for young children with autism: continuum-based behavioral models. *J Assoc Pers Sev Handicaps*. 1999;24:162–173.
52. Dawson G, Osterling J. Early intervention in autism: effectiveness and common elements of current approaches. In: Guralnick MJ, editor. *The Effectiveness of Early Intervention: Second Generation Research*. Baltimore (MD): Brookes; 1997:307–326.
53. Green G. Evaluating claims about treatments for autism. In: Maurice C, Green G, Luce SC, editors. *Behavioral Intervention for Young Children With Autism. A Manual for Parents and Professionals*. Austin (TX): PRO-ED; 1996:15–27.
54. Sallows GO, Graupner TD. Intensive behavioral treatment for children with autism: four-year outcome and predictors. *Am J Ment Retard*. 2005; 110:417–438.
55. Strauss K, Mancini F, Fava L; SPC Group. Parent inclusion in early intensive behavior interventions for young children with ASD: a synthesis of meta-analyses from 2009 to 2011. *Res Dev Disabil*. 2013; 34(9):2967–2985.
56. Berkowitz PB, Graziano AM. Training parents as behaviour therapist: a review. *Behav Res Ther*. 1972;10:297–317.
57. Vismara LA, Rogers SJ. Behavioral treatments in autism spectrum disorder: what do we know? *Annu Rev Clin Psychol*. 2010;6:447–468.
58. Reichow B, Servili C, Yasamy MT, Barbui C, Saxena S. Non-specialist psychosocial interventions for children and adolescents with intellectual disability or lower-functioning autism spectrum disorders: a systematic review. *PLoS Med*. 2013;10(12):e1001572. Epub 2013 Dec 17.
59. Oono IP, Honey EJ, McConachie H. Parent-mediated early intervention for young children with autism spectrum disorders (ASD). *Cochrane Database Syst Rev*. 2013;30(4):CD009774.
60. Fenske EC, Zalski S, Krantz PJ, McClannahan LE. Age at intervention and treatment outcome for autistic children in a comprehensive intervention program. *Anal Intervention Dev Disabil*. 1985;5:49–58.
61. Anderson SR, Campbell S, Cannon BO. The may center for early childhood education. In: Harris SL, Handleman JS, editors. *Preschool Education Programs for Children With Autism*. Austin (TX): PRO-ED; 1994:15–36.
62. Birnbrauer JS, Leach DJ. The Murdoch early intervention program after 2 years. *Behav Change*. 1993;10:63–74.
63. Lovaas OI. Behavioral treatment and normal educational and intellectual functioning in young autistic children. *J Consult Clin Psychol*. 1987;55:3–9.
64. Rogers SJ, Vismara L, Wagner AL, McCormick C, Young G, Ozonoff S. Autism treatment in the first year of life: a pilot study of infant start, a parent-implemented intervention for symptomatic infants. *J Autism Dev Disord*. 2014;44(12):2981–2995.
65. Harris SL, Handleman JS. Age and IQ at intake as predictors of placement for young children with autism: a four- to six-year follow-up. *J Autism Dev Disord*. 2000;30(2):137–142.
66. Ventola PE, Oosting D, Anderson LC, Pelphrey KA. Brain mechanisms of plasticity in response to treatments for core deficits in autism. *Prog Brain Res*. 2013;207:255–272.
67. Rogers SJ. Empirically supported comprehensive treatments for young children with autism. *J Clin Child Psychol*. 1998;27(2):168–179.
68. Schalock RL, Borthwick-Duffy SA, Bradley VJ, et al. *Intellectual Disability: Definition, Classification, and Systems of Supports*. Washington DC: American Association on Intellectual and Developmental Disorders; 2010.
69. Vivanti G, Barbaro J, Hudry K, Dissanayake C, Prior M. Intellectual development in autism spectrum disorders: new insights from longitudinal studies. *Front Hum Neurosci*. 2013;5(7):354.
70. Gabriels RL, Hill DE, Pierce RA, Rogers SJ, Wehner B. Predictors of treatment outcome in young children with autism: a retrospective study. *Autism*. 2001;5(4):407–429.
71. Sheinkopf SJ, Siegel B. Home-based behavioral treatment of young children with autism. *J Autism Dev Disord*. 1998;28(1):15–23.
72. Estes A, Olson E, Sullivan K, et al. Parenting-related stress and psychological distress in mothers of toddlers with autism spectrum disorders. *Brain Dev*. 2013;35(2):133–138.
73. Bebko JM, Konstantareas MM, Springer J. Parent and professional evaluations of family stress associated with characteristics of autism. *J Autism Dev Disord*. 1987;17(4):565–576.

74. Koegel RL, Koegel LK, Surratt AV. Language intervention and disruptive behavior in preschool children with autism. *J Autism Dev Disord*. 1992;22:141–153.
75. Herring S, Gray K, Taffe J, Tonge B, Sweeney D, Einfeld S. Behaviour and emotional problems in toddlers with pervasive developmental disorders and developmental delay: associations with parental mental health and family functioning. *J Intellect Disabil Res*. 2006;50:874–882.
76. Davis NO, Carter AS. Parenting stress in mothers and fathers of toddlers with autism spectrum disorders: associations with child characteristics. *J Autism Dev Disord*. 2008;38(7):1278–1291.
77. Wing L. The autistic spectrum. *Lancet*. 1997;350:1761–1767.
78. Kohane IS, McMurphy A, Weber G, et al. The co-morbidity burden of children and young adults with autism spectrum disorders. *PLoS One*. 2012;7(4):e33224. Epub 2012 Apr 12.
79. Munshi KR, Gonzalez-Heydrich J, Augenstein T, D'Angelo EJ. Evidence-based treatment approach to autism spectrum disorders. *Pediatr Ann*. 2011;40(11):569–574.

Supplementary materials

The comparison algorithms

In this section, we have briefly described the classic learning machines we compared. We have implemented the following learning machines using the WEKA software package (Waikato Environment for Knowledge Analysis, version 3.6.8, 1999–2012, an open source software tool developed for machine learning at the University of Waikato in New Zealand) and Semeion Software Suites (Rome, Italy; Buscema M, Supervised ANNs and Organisms, Semeion Software #12, version 23.0, 1999–2014).

Bayesian algorithms

The Bayesian algorithms are, obviously, based on Bayes' theorem, which states that given a set of events that partition an event space, any event dependent on event space enriches the knowledge of initial events by the equation:^{1,2}

$$P(E_i | A) = \frac{P(A | E_i)P(E_i)}{\sum_{j=1}^n P(A | E_j)P(E_j)} \quad (1)$$

The classifiers based on Bayesian networks (Bayes Net) represent the variables described by the formula in the equation¹ without special restrictions, whereas the naïve Bayesian networks (Naïve Bayes) are based on Bayes' formula with the assumption of stochastic independence between the variables. This drastic restriction of the domain of validity of the theorem makes this a high-performance classifier applicable to many practical problems.^{3–5}

The Naïve Bayes classifier used in this paper is according to the WEKA implementation.

Regression algorithms: logistic regression and multilayer perceptron

The logistic regression is a particular case of generalized linear regression applied in cases where the dependent variable “y” and its type are dichotomous.^{6,7}

The model is described by the function

$$\ln\left(\frac{p}{1-p}\right) = \beta_0 + \beta_1 x_1 + \dots + \beta_k x_k \quad (2)$$

with x_i independent variables and p is the probability that event y will occur.

As a generalization of the logistic regression model with a feed forward flow and totally interconnected, we have the multilayer preceptor model.⁸

The regression classifier and the multilayer perceptron classifier used in this paper follow the WEKA implementation.

Optimization algorithms: sequential minimal optimization and support vector machine

A support vector machine is a binary classifier that recognizes the hyperplane separating two different classes by maximizing the distance between the closest training examples.

Given a set of dual training

$$\{(x_i, y_j) | x_i \in \mathbb{R}^n, y_i \in \{-1, 1\}\}_{i=1}^n \quad (3)$$

we seek a solution for the equation

$$\max_{\alpha} \sum_{i=1}^n \alpha_i - \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n y_i y_j K(x_i, x_j) \alpha_i \alpha_j \quad (4)$$

in which

$$0 \leq \alpha_i \leq C; \quad i = 1; \quad \sum_{i=1}^n y_i \alpha_i = 0 \quad (5)$$

and where C = constant, $K(x_i, x_j)$ is the kernel function, and α_i represents Lagrange multipliers.

The sequential minimal optimization are iterative algorithms used to solve the optimization problem described for the support vector machine by decomposing it into a series of sub-problems, most small enough so that they can be solved analytically.^{9–12}

The sequential minimal optimization classifier used in this paper is according to the WEKA implementation.

Tree algorithms

Tree algorithms, or decision-making trees, rely on building a tree from the element's attributes (nodes) and the possible values that they can take (strings) until one arrives at the leaves representing the class of the instance. The path from the root node to a leaf node through the arch value determines the path that a particular instance must take to reach the membership class. The constructed tree attained from training datasets uses equations that determine the number of strings needed to be generated from a single node. Such decision trees can be used as binders.

J48

The J48 and the WEKA implementation of the C 4.5 algorithm was used to generate a decision tree of the kind developed by Ross Quinlan as an extension of the Iterative Dichotomiser 3 algorithm.¹³ A decision tree constructed in this way builds from the training data using the concept of entropy of a discrete random variable $X = \{x_1, \dots, x_n\}$

$$H(X) = - \sum p(x_i) \log p(x_i) \quad (6)$$

where $p(x_i)$ is the probability of the i th event.

Random trees, random forest

Random decision trees were introduced by Leo Breiman and Adele Cutler to treat both the problems of classification and regression. These are defined as a collection of decision trees called a forest.¹⁴ The random tree classifier takes in input feature vectors, the ranking for each tree in the forest, and assigns the class that had the largest number of recurrences.

J48 and random forest classifiers used in this paper are according to the WEKA implementation.

Rotation forest

Rotation forest¹⁵ draws upon the random forest idea. The base classifiers are also independently built decision trees, but in rotation forest, each tree is trained on the whole dataset in a rotated feature space. As the tree learning algorithm builds the classification regions using hyperplanes parallel to the feature axes, a small rotation of the axes, using principal component analysis, may lead to a very different tree.

Enhanced Back Propagation

Enhanced Back Propagation is an enhanced version of classic Back Propagation algorithm. The momentum is transformed in self-momentum in order to adapt the learning process to the local error condition of each network's node.¹⁶

Sine Net

Sine Net is characterized by the presence of a specific double nonlinear relationship on the connections between nodes. This characteristic has deep evident consequences on the properties of this network both on the computed function and behavior of this network during the learning phase.¹⁷⁻¹⁹

Instance-based learning algorithms

Instance-based learning algorithm is a sort of K-nearest neighbors classifier. It can select appropriate value of K based on cross-validation. It can also do distance weighting. The algorithm can work on numeric class, binary class, date class, nominal class, missing class values, and on the following types of attributes: date attributes, unary attributes, numeric attributes, nominal attributes, missing values, binary attributes, and empty nominal attributes.²⁰

References

1. Nielsen S, Nielsen TD. Adapting Bayes network structures to non-stationary domains. *Int J Approx Reason*. 2008;49:379–397.
2. Friedman N, Geiger D, Goldszmidt M. Bayesian network classifiers. *Mach Learn*. 1997;29:131–163.
3. Zhang H, Ling CX. Learnability of augmented Naive Bayes in nominal domains. In: Brodley CE, Danyluk AP, editors. *Proceedings of the Eighteenth International Conference on Machine Learning*. Burlington: Morgan Kaufmann. 2001:617–623.
4. John GH, Langley P. Estimating continuous distributions in Bayesian classifiers. In: *Proceedings of the eleventh conference on uncertainty in artificial intelligence*; 2005; Morgan Kaufmann Publishers, San Mateo.
5. Rish I. An empirical study of the naïve Bayes classifier. In: IBM research report, RC 22230, (W0111-014); 2001 New York.
6. Cessie S, van Houwelingen JC. Ridge estimators in logistic regression. *Appl Stat*. 1992;41:191–201.
7. Hosmer DW, Lemeshow S. *Applied Logistic Regression*. 2nd ed. New York: Wiley; 2000.
8. Rumelhart DE, Hinton GE, Williams RJ. Learning internal representations by error propagation. In: Rumelhart DE, McClelland JL, editors. Vol 1. *Parallel Distributed Processing*. Boston: the MIT Press; 1986:318–362.
9. Platt J. Fast training of support vector machines using sequential minimal optimization. In: Schoelkopf B, Burges CJC, and Smola AJ, editors. *Advances in Kernel Methods – Support Vector Learning*. Cambridge (MA): MIT Press; 1998.
10. Keerthi SS, Shevade SK, Bhattacharyya C, Murthy K. Improvements to Platt's SMO algorithm for SVM classifier design. *Neural Comput*. 2001; 13:637–649.
11. Keerthi SS, Gilbert EG. Convergence of a generalized SMO algorithm for SVM classifier design. *Mach Learn*. 2002;46:351–360.
12. Keeman V. *Learning and Soft Computing – Support Vector Machines, Neural Networks, Fuzzy Logic Systems*. Cambridge (MA): The MIT Press; 2001.
13. Quinlan JR. C4.5: Programs for Machine Learning, San Mateo: Morgan Kaufman, 2004.
14. Breiman L. Random forest. *Mach Learn*. 2001;45:5–32.
15. Rodriguez JJ, Ludmila I, Kuncheva C. Rotation forest: a new classifier ensemble method. *IEEE Trans Pattern Anal Mach Intell*. 2006; 28(10):1619–1630.
16. Arisawa R, Watada J. Enhanced back-propagation learning and its application to business evaluation. *Neural Networks*. 1994;1:155–160.
17. Buscema M, Terzi S, Breda M. Using sinusoidal modulated weights improve feed-forward neural network performances in classification and functional approximation problems. *WSEAS Trans Inform Sci Appl*. 2006;5(3):885–893.
18. Buscema M. Sine Net: an artificial neural network. Applicant Semeion Research Centre. Inventor M. Buscema, European Patent (Application n. 03425582.8 deposited 09-09-2003). USA Patent No US 7,788,196 B2 – Aug. 31, 2010. International Patent: Application PCT/EP2004/05189 deposited 08-28-2004.
19. Buscema M, Terzi S, Breda M. A Feed Forward Sine Based Neural Network for Functional Approximation of a Waste Incinerator Emissions: Proceedings of the 8th WSEAS Int Conference on Automatic Control, Modeling and Simulation, Prague, Czech Republic, March 12–14, 2006;276–280.
20. Aha D, Kibler D. Instance-based learning algorithms. *Machine Learning*. 1991;6:37–66.

Neuropsychiatric Disease and Treatment

Publish your work in this journal

Neuropsychiatric Disease and Treatment is an international, peer-reviewed journal of clinical therapeutics and pharmacology focusing on concise rapid reporting of clinical or pre-clinical studies on a range of neuropsychiatric and neurological disorders. This journal is indexed on PubMed Central, the 'PsycINFO' database and CAS,

Submit your manuscript here: <http://www.dovepress.com/neuropsychiatric-disease-and-treatment-journal>

Dovepress

and is the official journal of The International Neuropsychiatric Association (INA). The manuscript management system is completely online and includes a very quick and fair peer-review system, which is all easy to use. Visit <http://www.dovepress.com/testimonials.php> to read real quotes from published authors.