

A Study On The Impact Of Machine Learning Tools For Detecting Anxiety Disorders In Preschool Children

Anugraha K R¹ and Vineetha S²

¹ Student of Computer Science & Engineering , RIT ,Kottayam.

² Faculty of Computer Science & Engineering , RIT ,Kottayam.

E-mail: anugrahakr@gmail.com

Abstract. Nowadays anxiety disorders in preschool children is an interesting area for researchers. There are different types of anxiety disorders in which generalized anxiety disorder and separation anxiety disorders has high chance of occurrence among preschool children. Early identification of anxiety disorders in childhood can often prevent the persistence of these difficulties and appropriate treatments can improve the quality of their life. In light of the high prevalence of these disorders and the associated short and long-term impairment, there is a critical need for developing reliable and valid methods for measuring and monitoring the severity of these disorders and their symptoms. This survey discusses different studies in preschool anxiety disorders and different methods that are used as scales for measuring anxiety disorders. This paper also describes a machine learning technique using ADTree classifier for developing a screening tool for detecting anxiety disorders in preschool children. It was found that ADScore generated by ADTree can be used to measure the severity of disorder.

1. Introduction

Childhood plays an important role in structuring the social activities of a person. The social and emotional competencies are recognized as critical factors for children's success in later stages of adulthood [1]. Also the interaction of children among their primary caretakers and children set the stage for future academic and personal growth [2]. Even though a child is given utmost care, some may experience anxiety disorders due to some stress full events, which may develop into serious psychological disorders in their future.

Anxiety is normal and common in childhood. In most cases, anxiety in children is temporary and may be triggered by a despairing event. For example, a young child may experience separation anxiety when starting preschool or kindergarten. In some cases, however, this anxiety can be persistent and intense and can interfere with a child's daily routines and activities such as going to school, making friends, or sleeping. Such constant and serious anxieties that doesn't go away with reassurance and comfort, are called anxiety disorders.

According to Diagnostic and Statistical Manual of Mental Disorders, anxiety disorders include disorders that share features of excessive fear and anxiety and related behavioral disturbances. The most common types of anxiety disorders are Generalized anxiety disorder (GAD), Separation anxiety disorder (SAD),



Social Phobia and post-traumatic stress disorder. Early detection and diagnosis of such psychiatric disorders is crucial for their subsequent development and quality of life.

This paper discusses about the various anxiety disorders and their frequency in children. This paper also discusses about the different methods which are followed till date for finding the anxiety disorder.

2. Related works

Jessie C. Anderson et al [3] conducted study among preschoolers about anxiety disorders. In this study they found out that the most prevalent disorders were attention deficit, oppositional, and separation anxiety disorders(SAD), and the least prevalent were depression and social phobia. Conduct disorder, overanxious disorder, and simple phobia had intermediate prevalence. They were also able to find that the children identified at 11 years as having multiple disorders had a history of behavior problems since 5 years of age. Fifty-five percent of the disorders occurred in combination with one or more other disorders, and 45% as a single disorder. Orvaschel et al [4] also studied about the anxiety disorders among children. They reported that these anxiety syndromes were found in about 1-5% of children. They also pointed out that many of these anxiety disorders will be found in association with each other.

Lauren Franz et al. found that GAD, SAD and social phobia are the most common anxiety disorders found in preschool children. They conducted the study among reasonable representative samples of preschool aged children, using structured psychiatric assessments with known psychometric properties. In their study, they were able to find that about 1% of the children suffered from anxiety disorders. The range of prevalence estimates for generalized anxiety disorder reported in these studies is 0.6% to 6.5%, whereas prevalence estimates of separation anxiety disorder range from 0.3% to 5.4%, and social phobia range from 0.5% to 4.4%. They also found out that these anxiety disorders are highly comorbid at all ages. Comorbidity data from the preschool period remains limited with studies showing the co-occurrence between broad categories of emotional and behavioral disorders. Through this study they also reported that, comorbidity, even in the preschool years, is perhaps expected, as population-based twin studies indicate that the genetic liability for specific anxiety disorders partly overlaps.

The Duke Preschool Anxiety Study[9] aimed to do the following: establish prevalence rates for generalized anxiety disorder, separation anxiety disorder, and social phobia according to the DSM-IV criteria; detail patterns of comorbidity among generalized anxiety disorder, separation anxiety disorder, and social phobia; and examine the patterns of comorbidity among generalized anxiety disorder, separation anxiety disorder, and social phobia and other anxious or non-anxious disorders in preschool-aged children in pediatric primary care. The Duke Preschool Anxiety Study was done in three phases over a long period of time. The different phases include 1). The primary care questionnaire screening phase, 2). An in-home parent interview phase and 3). Lab based case control phase. From this study, they pointed out that among the preschoolers with anxiety disorders, only one fourth of them contains symptoms of more than one of the anxiety disorders. They also found that the Generalized Anxiety Disorder displayed the greatest degree of comorbidity with other anxiety disorders.

These anxiety disorders can be more commonly found in the early stages of life itself which may lead to serious psychological disorders if left undiagnosed. Appropriate measures should be found out for finding these disorders in the early stage itself. Some of the different practices or scales that are used for finding the anxiety disorders are discussed below.

2.1. Pediatric Anxiety Rating Scale (PARS)

The Pediatric Anxiety Rating Scale (PARS)[5] was developed as a clinically rated instrument for assessing both impairments and overall anxiety severity. It can be used to examine the course of anxiety symptoms and for judging the impact of treatments on the extremity and intervention of anxiety

symptoms in youths. In this work they created a checklist for screening the children and youths. The checklist was derived mostly from the *DSM-IV* criteria for anxiety disorders including social phobia (SoP), separation anxiety disorder (SAD), and generalized anxiety disorder (GAD) and by collecting details from a group of expert child and adolescent psychiatrists and psychologists at the participating sites of the Research Units on Pediatric Psychopharmacology (RUPP) who were asked to comment on: 1). Face validity, 2). Comprehensiveness, 3).Overlap with other constructs, and 4).Readability and clarity. This 50-item symptom checklist organizes items into the following categories: Social Interactions or Performance Situations (9 items), Separation (10 items), Generalized (8 items), Specific Phobia (4 items), Physical Signs and Symptoms (13 items), and other (6 items).

The PARS scale assessment is done by collecting information about symptoms from interviews with both child and parent(s).Endorsed symptoms are then collectively rated by the clinician on seven dimensions of severity, using a 6-point scale (0 for none, and 1–5 for minimal to extreme) for each dimension. These dimensions are (1) number of symptoms (none to more than 10), (2) frequency (none to several hours per day), (3) severity of distress associated with anxiety symptoms (none to extreme), (4) severity of physical symptoms (none to extreme), (5) avoidance (none to extreme), (6) interference at home (no interference to totally or almost totally unable to function at home), and (7) interference out of home (no interference to totally or almost totally unable to function out of home such as with peers or at school). The PARS Total Score is then calculated by adding the above scales of dimensions excluding number of symptoms and severity of physical symptoms. A score of 3 on each of these 5-point scales indicates a clinically significant level of severity, avoidance, or interference.

2.2 Multidimensional Anxiety Scale for Children (MASC)

Johns S. March and his colleagues developed the Multidimensional Anxiety Scale for Children (MASC)[6]. The MASC was developed to evaluate common anxiety symptoms in children across the elementary junior and senior high school age range. So when collecting the details for the anxiety symptom assessment procedure,they opt for not to assume anything about the normative clustering of pediatric anxiety symptoms other than to hypothesize that specific descriptors should fall within emotional, cognitive, physical or behavioral symptom domains .

Several steps were involved in the creation of MASC scale:

1. Available self-report anxiety scales covering general and specific symptom domains as well as the *DSM-III-R* criterion items were inspected again.
 2. Each item/question from these measures was then arranged into the one of the four symptom domains specified above.
 3. The number of items in each pool were reduced by:
 - Retaining items that can be easily understood, covered the desired age range, and closely reflected one of the four chosen anxiety dimensions.
 - Eliminating duplicates and rewording.
 4. Expert clinicians, members of an anxiety disorders support group, and lay non-experts classified the selected 60 items(15 per group) into the four selected domains. A Q sort procedure was used for this, which increased the item-content validity.
 5. Based on the selected experts' comments and the pattern of classification, a 41 item, 4-point Likert scale-having approximately 10 items perhypothesized symptom domain-was developed.
- During further study, they were able to find out that the MASC structure, which reflects the in vivo structure of pediatric anxiety symptoms, is invariant across gender and age and shows excellent internal reliability. Shared variance with scales sampling symptom domains of interest was highest for anxiety, intermediate for depression, and lowest for externalizing symptoms, indicating adequate convergent and divergent validity.

2.3. Screen for Child Anxiety Related Emotional Disorders (SCARED)

The Screen for Child Anxiety Related Emotional Disorders (SCARED)[7], is a self-report instrument that was developed as a screening tool for childhood anxiety disorders. Several steps were involved in the creation of final SCARED items. At first an 85 item questionnaire was generated by experienced clinicians, identifying symptoms of separation anxiety disorder (SAD), generalized anxiety disorder (GAD), social phobia, and school phobia (a simple phobia) based on *DSM-IV* classification of anxiety disorders. The questionnaire was then administered to children between the ages of 9 and 18 years and to their parents who attended the Child Mood and Anxiety Outpatient Clinic at Western Psychiatric Institute and Clinic in Pittsburgh. The severity of symptoms for 3 months was rated using a 0 to 2 point rating scale, with 0 meaning not true or hardly ever true, 1 meaning sometimes true, and 2 meaning true or often true and the collected data was used for preliminary analysis. Finally, a 38-item SCARED was derived after the factor analysis and deletion of items that overlapped with depressive symptoms.

2.4. Child behaviour checklist(CBCL)

The CBCL[11] is a standardized form which parents could use to describe their children's behavioral and emotional problems. The CBCL consists of a set of questions and for each question the parents could mark their response as 0 if the item is not true for their children, 1 if the item is somewhat or sometimes true and 2 if the item is very true or often true. There are two versions for CBCL form, the first version for children of age 2-3 and the second version for children of age 4-18. The CBCL/2-3 syndromes are designated in six areas: anxious/depressed, withdrawn, sleep problem, aggressive behavior and destructive behavior. The CBCL/4-18 consists of additional competence items to assess the child's activities, social relation and school functioning. The symptoms for these syndromes are rated by the parents based on their child's behavior for 2 months in CBCL/2-3 and for 6 months in CBCL/4-18.

The data obtained with the CBCL are summarized on a profile that displays the parent's rating of each item. The child's score on each syndrome is the sum of numbers that their parents marked on the individual items that comprise the syndrome. The profile could be used to analyze the child's standing on syndromes of problem that were derived from statistical analyses of CBCLs filled out for large number of clinically referred children. On the profile there are 2 broken lines which indicate a border line range between the normal and clinical ranges, the bottom line through 95% and top line through 98%. The score that are below the bottom broken line are considered normal and those that are above the top broken line are in clinical range. Scores between the broken line are high enough to be of concern but not high enough to be considered very deviant.

2.5. Preschool Age Psychiatric Assessment (PAPA)

Preschool Age Psychiatric Assessment (PAPA) [12] by Helen Link Egger is another effective screening tool for anxiety disorders. The PAPA is a parent interview for diagnosing psychiatric disorders in preschool children based on the parent version of the Child and Adolescent Psychiatric Assessment (CAPA). The trained PAPA interviewers interview the parent or other primary caregiver. A 3-month primary period is given to the parents for the assessment. There are 2 versions for PAPA. One is paper version and other is ePAPA. If the paper version is used then the individual PAPA variables are then entered into a database. If the ePAPA or electronic version of PAPA is used then the interview and coding are completed on a PC, making separate data entry unnecessary.

PAPA separately assesses the presence of the symptoms and the disabilities resulting from these symptoms. The symptoms are assessed mainly in 4 domains. They are: DSM-IV-TR and ICD-10 diagnostic criteria, Research Diagnostic criteria, and relevant behaviors and symptoms experienced by preschoolers and their families. 30 areas are covered while assessing the disabilities which includes the

child's relationship with his or her parents, other adults, siblings and peers as well as child's functioning in home, at school and out of the home.

One of the main advantages of PAPA is that, it combines the characteristics of both interviewer-based and respondent-based interview. Like respondent-based interviews, the PAPA uses a highly structured protocol, with required questions and probes. However, as trained interviewers are used, they could provide clear examples about behaviors or feelings relevant to the symptom, and have the symptom at a prespecified level of severity which helps the subjects understand the question clearly.

Egger et al. conducted the test-retest reliability of Preschool Age Psychiatric Assessment[12]. Three phases were used for performing the test, the screening phase, Test Re- Test phase (TRT) and data analysis phase. In the screening phase equal number of parents of boys and girls were selected based on some selection criteria. In the TRT phase, selected parents were subjected to interviews. The interviewers were blind to the parent's screen status. After the first interview a second interview was scheduled within a period of about 1 week. The second interview was by an interviewer who was blind to the results of the first. Cohen's k method was used to assess agreement on categorical variables, and the intra class correlation coefficient (ICC) was used to assess agreement between syndrome scale scores. Computed weighted reliability statistics from the whole sample was used to produce unbiased estimates of reliability for pediatric primary care. These weights were inversely proportional to the probability of selection into the test-retest sample. They also computed non weighted estimates for the screen high group alone to approximate expected reliability in psychiatric clinic samples. They achieved diagnostic reliability similar to those obtained from interviews with k ranging from 0.36 to 0.79. They also achieved good reliability in the assessment of disability and SED which proves that PAPA could be helpful for both researchers and clinicians. As the study included assessment of children of different ages and from different backgrounds, they suggest that PAPA could be used with various population of children.

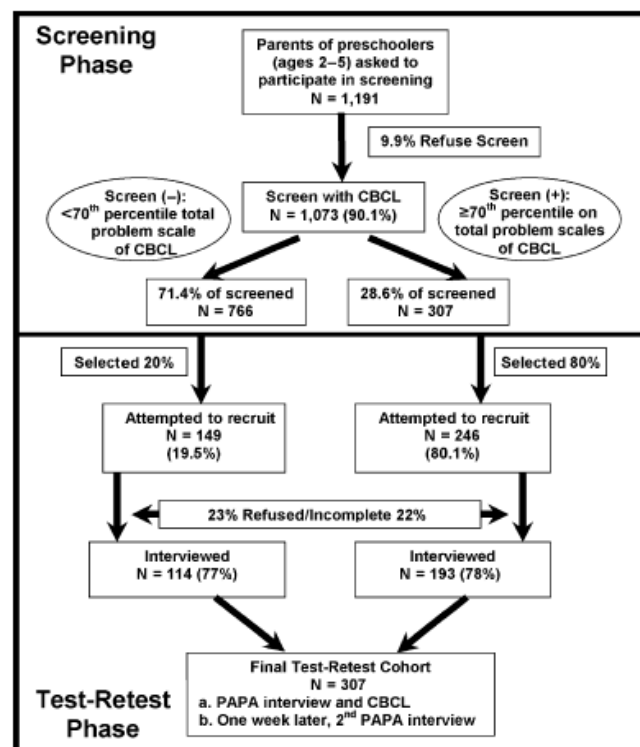


Figure 2.5.1 : Design of PAPA test-retest study[12]

Thus the assessment of psychopathology in childhood is commonly carried out by performing long comprehensive interviews such as Preschool Age Psychiatric Assessment (PAPA). Unfortunately, the time required to complete a full interview is too long to apply it at the scale of the actual population at risk, and most of the population goes undiagnosed or is diagnosed significantly later than desired. So numerous works are being done to find the relationship among the different factors/items and combine them, thereby reducing time required for completing these different assessment scales and methods

One such notable work was done by Birhameret. al [8]. In this work they tried to reduce the number of items in SCARED and provide same reliability as that of the full version SCARED. Factor analysis was conducted by administering the SCARED to a group of children and their parents. In the preliminary analysis, they found that the social factor did not discriminate well between patients with social phobia and other anxiety disorders. So 3 items were added to the 38-item version for the conducting the test. Based on the factor analysis on the 41 item version, 5 factors were obtained: panic/somatic, generalized anxiety, separation anxiety, social phobia, and school phobia. The 5 factor SCARED demonstrated good internal consistency and discriminant validity.

The recent study about preschool anxiety disorders was held by Kimberly Carpenter [13] and his colleagues. In this work, they aim to learn from unique and very rich previously collected PAPA examples the inter-correlations between different questions in order to provide a reliable risk analysis in the form of a much shorter interview. They use for this purpose the alternating decision trees algorithm, which combines decision trees with boosting to produce small and interpretable decision rules. Rather than a binary prediction, the algorithm provides a measure of confidence in the classification outcome. This is highly desirable from a clinical perspective, where it is preferable to abstain a decision on the low-confidence cases and recommend further screening. In order to prevent over-fitting, they propose to use network inference analysis to predefine a set of candidate question with consistent high correlation with the diagnosis.

Among the different measurement techniques for anxiety disorders, PAPA is more reliable and most widely used. PAPA include a large number of questions and therefore it may take hours to complete one PAPA interview. So it is important to reduce the number of questions to detect anxiety disorders without compromising the reliability of the test. Even though works are done regarding reducing the complexity of the PAPA assessment, a more reliable and shorter version of PAPA to assess the most commonly found anxiety disorders would be highly beneficial.

Since the dataset provided by PAPA is high dimensional a screening tool is very essential to detect the disorder. After PAPA interview the collected data must be clinically analyzed to detect the disorder. This whole procedure is very time consuming. Machine learning technique using ADTree classifier is found to be very effective to develop screening tool to detect anxiety disorder[14]. In order to reduce the complexity of constructing ADTree different machine learning techniques can be used to select most informative features from PAPA dataset .

3. ADTree Classifier

An alternating decision tree(ADTree) is a decision tree consists of alternating layers of prediction nodes and splitter nodes. The values in a root node represent the initial probability for assigning the target class according to the training dataset. Alternating decision tree classifiers are then built according to a particular structure using boosting wherein simple rules are successively added to the alternating decision tree classifier until the unit classifier of the tree exhibits satisfactory performance. Since boosting iteration adds three nodes (one splitter node and two prediction nodes) to the tree, more boosting iterations will result in larger and potentially more accurate trees.

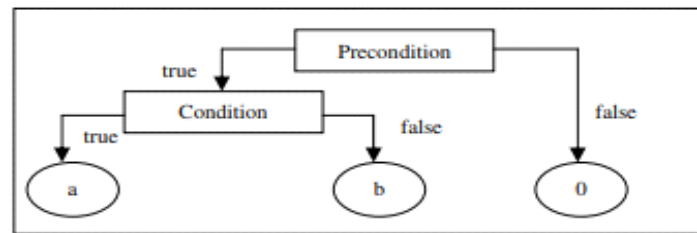


Figure 3.1: Graphical representation of single decision rule of ADTree[14]

Unlike original decision trees: an instance is mapped into a path along the tree from the root to one of the leaves and output is the label of the leaf, the classification result of an alternating decision tree became the sign of the sum of the predictions along the multi-path associated with the given instances. When some feature values are unknown, the alternating decision tree algorithm only considers the reachable decision nodes. Also, the alternating decision tree algorithm has shown competitive performances and has produced smaller and intuitive classification rules than general decision tree algorithms. The figure below is an example of ADTree.

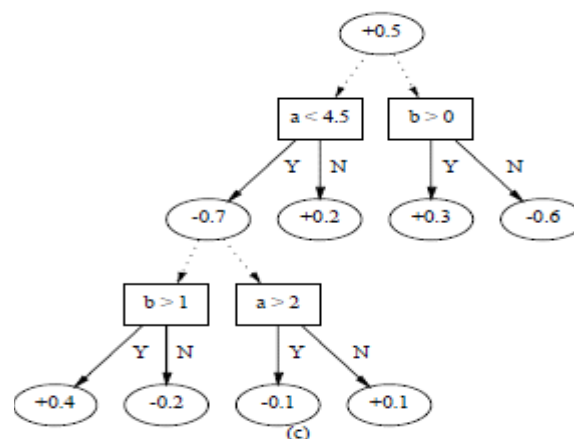


Figure 3.2 : Simple ADTree[14]

4. Conclusion

In this paper different methods such as PARS, MASC, SCARED, CBCL and PAPA for measuring anxiety disorders are analyzed. From that PAPA is found to be reliable and valid structured diagnostic parent report interview for detecting anxiety disorders. But PAPA is very time consuming as it involve a long set of questions. So the advantage of PAPA assessment can be utilized by extracting the essential questions that would identify the SAD and GAD. The most recent approach to develop a screening tool for detecting GAD and SAD is a machine learning based method. . ADTree classifier can be used to develop a screening tool for detecting anxiety disorder. Instead of making a binary decision, ADTree classifier produces an ADScore, which can be used to calculate risk score for GAD and SAD. Risk score value lies between zero and one. A score near to zero indicates less chance of disorder and score near to one indicates high chance of disorder.

5. References

- [1] Kagan, S.L., Moore, E. and Bredekamp, S. eds., 1998. *Reconsidering children's early development and learning toward common views and vocabulary*: National Education Goals Panel. DIANE Publishing..
- [2] Denham, S.A., 2006. *Social-emotional competence as support for school readiness: What is it and how do we assess it?*. Early education and development, 17(1), pp.57-89.
- [3] Anderson, J.C., Williams, S., McGee, R. and Silva, P.A., 1987. *DSM-III disorders in preadolescent children: Prevalence in a large sample from the general population*. Archives of general psychiatry, 44(1), pp.69-76.
- [4] Orvaschel, H. and Weissman, M.M., 1986. *Epidemiology of anxiety disorders in children: A review*. Anxiety disorders of childhood, pp.58-72.
- [5] Research Units on Pediatric Psychopharmacology Anxiety Study Group, 2002. *The pediatric anxiety rating scale (PARS): Development and psychometric properties*. Journal of the American Academy of Child & Adolescent Psychiatry ,1061-1069.
- [6] March, J.S., Parker, J.D., Sullivan, K., Stallings, P. and Conners, C.K., 1997. *The Multidimensional Anxiety Scale for Children (MASC): factor structure, reliability, and validity*. Journal of the American academy of child & adolescent psychiatry, 36(4), pp.554-565.
- [7] Birmaher, B., Khetarpal, S., Brent, D., Cully, M., Balach, L., Kaufman, J. and Neer, S.M., 1997. *The screen for child anxiety related emotional disorders (SCARED): Scale construction and psychometric characteristics*. Journal of the American Academy of Child & Adolescent Psychiatry, 36(4), pp.545-553.
- [8] Birmaher, Boris, et al. 1999. *Psychometric properties of the Screen for Child Anxiety Related Emotional Disorders (SCARED): a replication study*. Journal of the American Academy of Child & Adolescent Psychiatry 38.10 , 1230-1236.
- [9] Franz, L., Angold, A., Copeland, W., Costello, E.J., Towe-Goodman, N. and Egger, H., 2013. *Preschool anxiety disorders in pediatric primary care: prevalence and comorbidity*. Journal of the American Academy of Child & Adolescent Psychiatry, 52(12), pp.1294-1303..
- [10] Biederman, J., Wozniak, J., Kiely, K., Ablon, S., Faraone, S., Mick, E., Mundy, E. and Kraus, I., 1995. *CBCL clinical scales discriminate prepubertal children with structured interview—derived diagnosis of mania from those with ADHD*. Journal of the American Academy of Child & Adolescent Psychiatry, 34(4), pp.464-471.
- [11] Egger, H.L. and Angold, A., 2004. *The Preschool Age Psychiatric Assessment (PAPA): A structured parent interview for diagnosing psychiatric disorders in preschool children*. Handbook of infant, toddler, and preschool mental health assessment, pp.223-243.
- [12] Egger, H.L. and Angold, A., 2006. *Common emotional and behavioral disorders in preschool children: presentation, nosology, and epidemiology*. Journal of child psychology and psychiatry, 47(3-4), pp.313-337.
- [13] Carpenter, K., Sprechmann, P., Fiori, M., Calderbank, R., Egger, H. and Sapiro, G., 2014, May. *Questionnaire simplification for fast risk analysis of children's mental health*. In Acoustics, Speech and Signal Processing (ICASSP), 2014 IEEE International Conference on (pp. 6009-6013). IEEE.
- [14] Sok, H.K., Chowdhury, M.S., Ooi, M.P.L., Kuang, Y.C. and Demidenko, S., 2013, May. *Using the ADTree for feature reduction through knowledge discovery*. In Instrumentation and Measurement Technology Conference (I2MTC), 2013 IEEE International (pp. 1040-1044). IEEE.