An Explainable Genetic Programming Approach to Safely Predict Cyberbullying Occurrence in Ireland.

Abstract

Cyberbullying is growing a growing problem in Ireland, with reported rates of occurrence growing every year for both primary and secondary school students. We have collected survey data from primary school children across the country and asked them their beliefs about internet safety, their opinion of their own knowledge of the internet, as well as their actions online. This survey data, collected over 9 years, represents by far the largest dataset on cyberbullying ever collected and analysed in Ireland. We use this dataset to build an explainable machine learning classifier called a Fuzzy Pattern Tree. Fuzzy Pattern Tree classifiers achieve close to state-of-the-art results, attaining mean test accuracy of 84.3%, while allowing their internal workings to be examined. Examining the logic of the models ensures both their safe deployment and allows for effective interventions and corrections in behaviour to help children avoid experiencing cyberbullying. Our models show that increased awareness from parents about the apps their children use, as well as their social media activity are important to avoid cyberbullying. The Fuzzy Pattern Tree models also point towards smartphone usage as a major risk factor for cyberbullying.

Keywords

Cyberbullying, Genetic Programming, XAI, Fuzzy Logic

1. Introduction

Cyberbullying occurs online when using digital devices such as smartphones, computers, and tablets. It involves using technology to harass, threaten, or embarrass someone, often via social media or messaging apps. The occurrence and rates of cyberbullying differ greatly across regions, genders and ages for a number of different reasons including rate of technology usage, communication skills and membership of minority groups [1].

There are many negative effects on mental well-being that are linked with cyberbullying. These include psychological distress, decreased life satisfaction and even suicidal ideation [2]. Given that experiencing cyberbullying, particularly via social media, has been linked with these serious mental health problems, researchers and health care workers have explored ways to prevent, mitigate or intervene in these situations. The available data suggest that holistic programs that include a close and coordinated collaboration between schools, social welfare services and parents are needed to build programs to prevent and eliminate bullying and cyberbullying [3]. In particular, when dealing with adolescents and online bullying evidence suggests that active parental involvement and monitoring of social media use can be an effective solution and reduce negative outcomes for the victims [4]. This may not happen, however, as children may fear that their devices will be taken away if they report this bullying and will instead suffer silently. It is therefore critical for parents or other guardian figures to be aware of the signs of cyberbullying and closely monitor their children's activities.

Machine learning (ML) offers a solution to this problem as it may be able to automatically predict if a child is at risk for cyberbullying and allow for intervention to minimise, or remove completely, the harm done. In particular, an explainable ML, or explanable AI (XAI), solution may highlight certain actions or routines which a child is engaging in that particularly puts them at risk for cyberbullying and allows parents or teachers to effectively step in and correct such behavior.

XAI aims to create interpretable models and methods that can somehow explain themselves without, or with minimal, impact on performance. An XAI method which has shown strong performance and explainability are Fuzzy Pattern Trees (FPTs) [5]. Based on fuzzy set theory, an FPT is a hierarchical tree structure, not a rule list. Due to it's use of fuzzy operators (i.e linguistic labels) it is more easily interpretable. It has been shown that this interpretability, naturally, is contingent on the trees not being excessively large [6].

Zeeko Education[2] have been surveying children in Ireland for over 9 years and have collected over 100,000 responses from children about their attitudes towards internet safety, their behaviours online and their experience with cyberbullying. These surveys are by far the largest collection information about Irish children's online habits.

This paper uses this state-of-the-art dataset and a powerful XAI technique, FPTs, to build an explainable ML model to predict if primary school children are at risk of being cyberbullied.

Section 2 reviews the background to this research, including cyberbullying, the issues it generates, potential solutions and it's prevalence, particularly within Ireland. It also discusses Grammatical Evolution and FPTs, the models we use to predict cyberbullying. Section 3 details the Zeeko Internet Safety surveys, the questions asked, summary of the data and describes the surveys responses in more detail. It also presents the experimental set-up which was used, describes the various classifiers we benchmark our approach against and details the parameters used for each classifier. Section 4 presents the main results of the experiments described in 3. Finally, Section 5 summarises the research and discusses future work suitable for investigation.

2. Background

2.1. Cyberbullying in Ireland

Many reviews and surveys have been conducted examining various areas of both traditional bullying and cyberbullying in Ireland. In Ireland, as in many countries, cyberbullying has been shown to cause serious emotional and psychological harm, especially to young people. The impacts of cyberbullying differ from other forms of bullying in several ways. Cyberbullying can happen anytime, unlike traditional bullying, which is often limited to specific locations. Victims feel they can't escape as it follows them through their devices. Online bullies can remain anonymous, making the bullying harder to stop and increasing the victim's fear and helplessness. Cyberbullying can reach a wider audience quickly, making the humiliation more public. Harmful digital content can be difficult to erase, prolonging the victim's distress. Cyberbullying can lead to more intense psychological harm due to its constant nature, including heightened isolation and anxiety. These factors make cyberbullying particularly harmful, requiring specific prevention and intervention strategies.

While the rates of cyberbullying are lower than traditional bullying in schools, both in Ireland and worldwide, there has been a sharp increase in prevalence. A report from 2017 stated the prevalence of cyberbullying in the island of Ireland at 5.2% in primary schools and 3.9% in post-primary schools [7]. This has dramatically increased according to a 2023 survey which found that 25% of children aged 8-12 reported as suffering from some form of cyberbullying, rising to 40% for those aged 12-16 [8]. Another survey of talented adolescents in Ireland echoed these findings, reporting that just over 31% of the surveyed students had been a victim of cyberbullying, with 18.5% experiencing cyberbullying in the past 3 months.

This sharp increase has come in spite of large efforts by both governmental and non-governmental bodies to curb online abuse. In particular, Ireland has made some legal efforts to try criminalise cyberbullying, specifically "Coco's Law" [9]. Ireland is an outlier within the Member States of the European Union, however, as cyberbullying is scarcely regulated by common laws. In particular, there is no official EU law aimed at criminalising online harassment, victimisation or bullying.

Reducing cyberbullying for children in Ireland requires a combination of education and technological solutions. Educational strategies include teaching digital literacy in age-appropriate ways, fostering empathy, and addressing online privacy and bullying in both primary and secondary schools. Training for teachers and parents is also crucial to identify and address cyberbullying. Incorporating wellbeing into the curriculum can help students cope with its emotional impact. Youth mental health practitioners

[2] https://zeeko.ie/

have reported that more training and resources are needed for Child and Adolescent Mental Health Service staff and caregivers to effectively combat cyberbullying [10].

Technological solutions also require investigation and is the motivation for this paper. The focus of this technology should be on empowering students to navigate the digital world safely and responsibly. An automatic, personalised risk and behaviour ML model would allow for prompt and effective interventions to be made. This model would need to be explainable in order to deploy it safely and in order to identify the key variables the model uses and allow teacher and parents to create actionable plans for the children [11]. An XAI approach would also be required in order to comply with the recently introduced EU AI Act [12].

2.2. GE

Grammatical Evolution (GE) [13] is an evolutionary computation (EC) search technique which uses a grammar, generally, a context-free grammar (CFG) written in Backus-Naur form (BNF), to find syntactically correct executable programs which solve a given problem. As with many other evolutionary algorithms, GE's inspiration comes from observing nature, specifically genetics. GE creates programs (which can be trees, circuits, rules) by mapping an integer string into the final structure wanted using a grammar. A key difference between GE and other EC methods is that the evolutionary operators of mutation and crossover are performed on the string, not on the output structure which is tested. This separation between the search space and the program space has seen GE achieve success in a wide variety of domains, including digital circuit design [14], automatic test case generation [15] and Neural Network Optimisation [16].

Successfully separating the search space and program space is one of GE's great innovations. However, this comes as a cost as this separation leads to a disruptive effect known as *ripple*, also known as *ripple effects* [17]. Simply put, minor changes to the integer string of an individual, particularly in the first few digits, may have drastic effects on the resulting program. The program of a child solution may be almost entirely different from its parent despite there being very little variation, perhaps only one integer difference, in their respective strings. This can occur with both evolutionary operators, crossover and mutation.

To alleviate these concerns, the FPTs in this paper are evolved using Structured Grammatical Evolution (SGE) [18]. SGE overcomes the poor locality of GE and limits the ripple by altering the construction of the integer lists. In standard GE, a single list of numbers are used, left to right, to map to an individual to it's final output. In SGE, a set of lists is used, each list corresponding to a unique part, called a non-terminal, of the grammar. When that non-terminal is selected, the list corresponding to that non-terminal is used to complete the mapping and not the next number on the list, which is used in GE. This ensures that any change to a integer is confined to that non-terminal and that a crossover or mutation does not "ripple" throughout the solution.

2.3. FPTs

An FPT is a hierarchical ML model with has tree structure. It's internal nodes consist of fuzzy logical operators and fuzzy arithmetic operators, while it's leaf nodes are fuzzified input variables and constants. FPTs were first introduced, independent of each other, by [19] and [20], who called this type of model Fuzzy Operator Trees. An FPT model is closely related to other fuzzy logic model classes, including fuzzy rule-based systems (FRBS), and fuzzy decision trees (FDT).

FPTs, a white box ML method, which use evolutionary computation to optimise their structure have been shown to be competitive with, and sometimes outperform, black box methods [21]. Crucially, FPTs have been shown to allow users gain an understanding of their internal logic [22]. gate.ucd.ie

To perform classification using FPTs a set of FPTs is needed, one for each class that exists in the problem, the classifier decision occurs in favor of the tree (class) that has the highest output value for that instance. These FPTs serve as the logical description of the class and grants a more precise interpretation of what the model is doing and grants insight into how the problem is being solved.

In our experiments we evolves one, large solution and treat the subtrees of this solution as its FPTs, as seen in Figure 1. The FPT which yields the largest output for an individual, is declared the winner, and that individual is designated as belonging to that class. This is illustrated in Figure 2. The root node of the tree is responsible for this process. Representing each FPT as subtrees of one large solution combined with SGE's inbuilt separation between search space and program space leads to another major advantage our representation experiences. No special or protected operators are needed for crossover or mutation. A simple grammar augmentation is all that is needed to tackle different problem specifications.

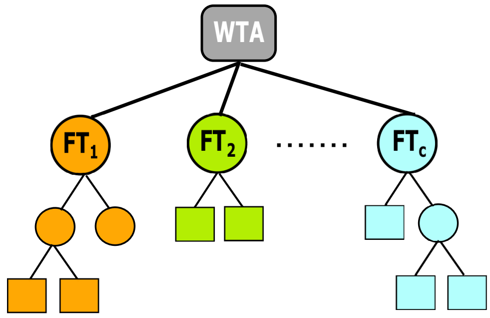

Figure 1: Pictorial representation of a multi-classifier evolved by SGE, where FT_c is the fuzzy tree for each available class, and at the root the winner take all (WTA).

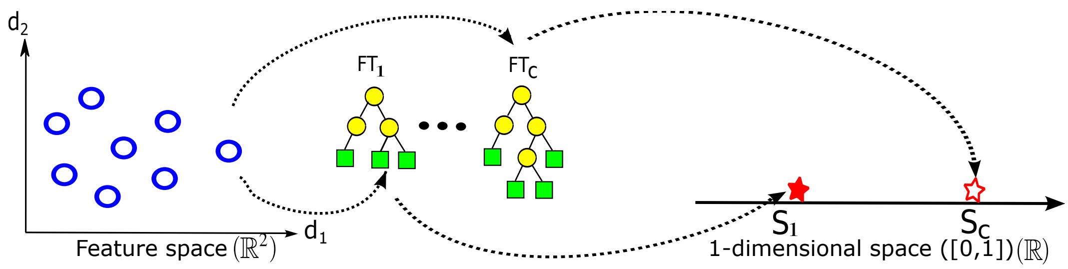

Figure 2: Graphical depiction of the mapping process from the feature space to a 1-dimensional space [0,1] using a set of fuzzy trees FT_1 to FT_c.

3. Experimental Setup

3.1. Survey and Dataset

The dataset used to train the FPTs and other ML models was gathered via surveys undertaken by children during an Internet Safety Seminar run by Zeeko Education. The survey's took place between 2016 and 2024 in primary schools in the Republic of Ireland. The questions asked in the survey which were used in our training data are described in Table 1.

A total of 79,260 surveys from primary school students were collected. After cleaning and removal of incomplete or erroneous surveys, the final dataset used consisted of 67,387 surveys. This is, by far, the largest dataset ever collected around the topic of cyberbullying in Ireland.

According to the dataset, the rate of experiencing cyberbullying in schools in Ireland is 15.6% (10,480 replied 'Yes' when asked if they have ever been cyberbullied). This is slightly below the rates reported in Section 2.1, but does reflect that cyberbullying has increased dramatically in the past decade. The gender balance of the dataset is slightly weighted towards females than males, 34,941 and 32,106 respectively.

The majority of the children surveyed were in senior classes, with 16,531 children in 6th class, 16,010 in 5th class and 14,999 in 4th class. 12,787 of the children were in 3rd class, followed by 6,609 and 1,563 in 2nd and 1st class, respectively.

Table 1
Questions Asked in Internet Safety Survey

Question	Possible Answer
What is your Gender?	Male, Female, Other
What devices do you use to access the internet, play games, use apps, etc?	Smartphone, Laptop, etc,
Do you think you know more than our parents about:	
Apps?	Yes/No
Online Gaming	Yes/No
Social Media	Yes/No
Internet in General	Yes/No
How much screen time do you usually get:	
On Weekdays?	0 hours, Less 1 hour,..., more 5 hours
On Weekends?	0 hours, Less 1 hour,..., more 5 hours
How serious are the following:	
Spending too long online	Not Serious At All,.., Very Serious
Cyberbullying	Not Serious At All,.., Very Serious
Talking to a person you met first online	Not Serious At All,.., Very Serious
To be careful with the posts photos and videos you put online	Not Serious At All,.., Very Serious
Have you ever:	
Spoken or chatted to a stranger online?	Yes/No
Played with or against a stranger online?	Yes/No
Played an over 18's game?	Yes/No
Have you ever been cyberbullied?	Yes/No

35,070 primary school students use a smartphone to access the internet, 42,081 use Tablets, 25,424 use Laptops/Desktops, and 30,120 use a Games Console.

A majority of children believe that they know more about Apps and Gaming than their parents, 40,195 and 47,904 but do not feel the same way about Social Media or the Internet in general, 21,061 and 26,015. 20,080 children said they have chatted to a stranger online and 37,081 have played with a stranger online. 19,426 admitted to playing an over 18's game.

Most children believe that spending too long online is kind of serious or serious, 28,306 and 20,608 respectively. 9,654 believe it is very serious while only 8,819 believe it is not serious at all. The vast amount of children, 51,1776, view cyberbullying as very serious compared to 4,440 saying it is not serious at all. 8,384 said it was serious and 2,787 said it was kind of serious.

27,844 students think that talking to a person you met online is very serious and 19,192 believe it is serious. 13,149 believe it is kind of serious while 7,202 responded that it is not serious at all. Finally, 36,196 student believe being careful when posting photos/videos online is very serious and 18,477 believe it is serious. Very few think it is kind of serious or not serious at all, 8,115 and 4,599 respectively.

Just over 17.5% (11,916) of children get over 5 hours screen time per day on weekdays. This almost doubles to 30.2% (20,333) on weekends. 12,932 of students spend less than 1 hour online during the week, falling to 5,838 on the weekend. 19,317 spend 1-2 hours online a day during the week and 12,197 spend 2-3 hours. On the weekend these figures are 14,499 and 15,291 respectively. 4,603 children responded that they have no screen time during the week, plummeting to 941 during the weekend.

3.2. Parameters

The full SGE experimental parameters are seen in Table 2. The maximum tree depth is set to 6 and reflects the depth at which FPTs have previously been shown to lose their interpretability [22]. We use sensible initialisation to create the population of solutions [23] and used Root Mean Square Error (RMSE) to guide the search [24].

We compare the results of the FPTs with 4 other ML methods; Logistic Regression, Random Forest [25], Support Vector Machines [26] and XGBoost [27]. The hyper-parameters for each of these methods underwent a simple grid-search optimisation prior to execution. 30 independent runs of each model were performed, with the dataset randomly split 75%/25% for training and test at the beginning of each run.

Table 2
List of the main parameters used to run SGE

Parameter	Value	Parameter	Value
Total Generations	50	Population	500
Elitism	5	Selection	Tournament (3)
Crossover	0.9	Mutation	0.1
Max Tree Depth	5	Min Tree Depth	2
Initialisation	Sensible	Fitness Function	RMSE

The following operators are used within the FPTs, where a and b are the inputs to the operator:

$$WTA = IF\{\}()..ELSE() \tag{1}$$
$$MAX = max(a, b) \tag{2}$$
$$MIN = min(a, b) \tag{3}$$
$$WA(k) = ka + (1-k)b \tag{4}$$
$$OWA(k) = k \cdot max(a, b) + (1-k)min(a, b) \tag{5}$$
$$CONCENTRATE = a^2 \tag{6}$$
$$DILATE = a^{\frac{1}{2}} \tag{7}$$
$$COMPLEMENT = 1 - a \tag{8}$$

where WTA, WA & OWA denote Winner Takes All, Weighted Average and Ordered Weighted Average, respectively.

The binary classification grammar used in experiments can be seen in Figure 3. The WTA node contains two $< exp >$ non-terminals which need to be expanded. These will be the FPTs for each class when they are fully expanded. Two FPTs are required for binary classification.

$$\begin{aligned}
< start >&::= WTA(< exp >, < exp >)\\
< exp >&::= max(< exp >, < exp >) \mid\\
&\quad min(< exp >, < exp >) \mid\\
&\quad WA(< const >, < exp >, < exp >) \mid\\
&\quad OWA(< const >, < exp >, < exp >) \mid\\
&\quad concentrate(< exp >) \mid\\
&\quad dilation(< exp >) \mid\\
&\quad complement(< exp >) \mid\\
&\quad x_1 \mid x_2 \mid x_3 \mid ...\\
< const >&::= 0. < digit >< digit >< digit >\\
< digit >&::= 0 \mid 1 \mid 2 \mid ....
\end{aligned}$$

Figure 3: Grammar used to evolve a Fuzzy Pattern Tree for a binary dataset. The WTA node can be augmented by adding extra $< exp >$ to include as many subtrees as necessary, making it a multi-class grammar.

3.3. Fitness function

The fitness functions used during the evolution of the FPTs, RMSE, is shown below in Eq. 9. As well as classification accuracy, we also report the Matthews correlation coefficient (MCC) for each classifier, Eq. 10.

$$\mathrm{FF}_{\mathrm{RMSE}} = \sqrt{\sum_{i=1}^{n} \frac{(\hat{y}_i - y_i)^2}{n}} \tag{9}$$

$$MCC = \left(\frac{TP * TN - FP * FN}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}} \right)^2 \tag{10}$$

4. Results and Discussion

4.1. Experimental Results

The full results of the experimentation can be seen in Table 3. It can be seen that Random Forest attained the best performance, averaging 85.7% across the 30 runs. It also found the best model found at 86.1%. Slightly behind this were SVM and FPTs, achieving an average accuracy of 84.4% and 84.3% respectively. Somewhat surprisingly, XGBoost only obtained a mean accuracy score of 79.5% and a best model of 82.1%. Performing far worse than both Random Forest and SVM's, as well as being almost 5% worse on average than FPT's, may be due to the nature of the data being mostly categorical. Another reason may be the experimental setup, while some hyper-parameter optimisation was performed for all the methods it was not exhaustive and there is scope for further improvement. The worst ML method considered was seen to be Logistic Regression, over fitting badly and only finding mean accuracy of 67.7%. While not unexpected to be the worst performing method, the large gulf in performance is remarkable. As with XGBoost, Logistic Regression may see an increase in performance with regularisation.

As well as accuracy, we report the MCC of each classifier. The MCC suggests the classifiers are not as powerful as the accuracy measures may suggest. The dataset has moderate, but not severe, imbalance and the MCC shows that all methods are overfitting and leveraging this imbalance to some extent. Random Forest has the best MCC at 0.271, followed by SVM, 0.266, and FPTs, 0.262. Despite having larger accuracy, XGBoost has a lower MCC than Logistic Regression, 0.258 and 0.259 respectively.

Echoing results seen previously, FPT's are able to attain close performance with other, black-box classification methods. However, those previous results were on datasets much smaller than the Cyberbullying dataset used here. Our results further reinforce that FPTs as a strong white-box, classifier.

Table 3
Test classification performance comparison of each model, showing the classification accuracy and MCC on the test data for the best solution found averaged across 30 runs. The standard deviation for both are shown in brackets. The final column, Best, contains the test accuracy from the best model from the 30 runs.

Method	Accuracy	MCC	Best
Logistic Regression	67.7% (0.3%)	0.259 (0.007)	69.8%
Random Forest	85.7% (0.2%)	0.271 (0.009)	86.1%
SVM	84.4% (0.1%)	0.266 (0.002)	84.9%
XGBoost	79.5% (0.4%)	0.258 (0.009)	82.1%
FPT	84.3% (0.2%)	0.262 (0.001)	84.5%

4.2. Interpreting the Models

As interpretability is a key concern in this study, the FPT's were next examined. The best models found (as one FPT is needed for each class, in this case two) are shown in Figure 4 and Figure 5. Figure 4 shows the FPT for predicting the occurrence of cyberbullying based on the survey responses and Figure 5 shows the FPT for predicting the absence of cyberbullying.

We can observe from Figure 4 that spending very little time online during the week but large amounts over the weekend increases the chance of experiencing cyberbullying. The model predicts that a child is at risk when they are careful with what they post online and believe that spending too much time online is a serious problem, top right subtrees, but they do not believe that talking to strangers they met online is a serious issue (bottom left). They also feel that are more informed that their parents, the bottom left subtree showing they believe they know more about Apps and Social Media than their parents. This shows the danger that having a partial but not complete knowledge of internet safety can have, as this limited knowledge may make them reckless in other areas as they incorrectly believe they are acting carefully (being careful with photos/videos but still talking with strangers). The model may also point towards parents being more vigilant of their children's online behaviour during the week, but not on weekends.

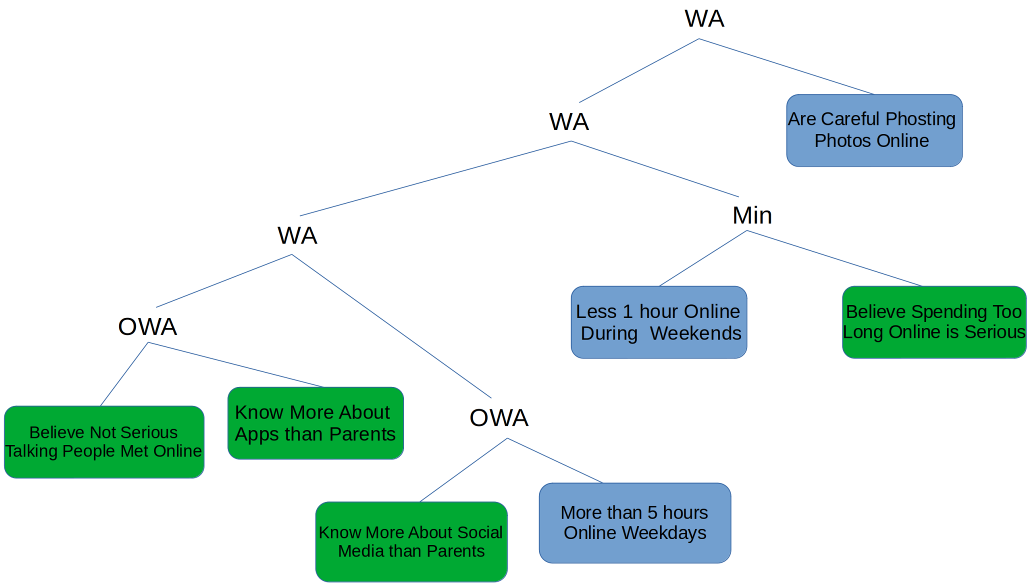

Figure 4: FPT used to Predict Cyberbullying. Green boxes show Attitudes while Blue boxes show behaviours. This tree shows that children that believe that they know more about Apps and Social Media than their parents and that it is not serious to talk to strangers they met online have a higher chance of being cyberbullied, as shown in the bottom left subtrees. Interestingly, spending large amounts of time online during the week but not at the weekends increases cyberbullying risk. Perhaps surprising, being careful with posting photos and believing that spending too long online is a serious problem also increase cyberbulyying risk.

Figure 5 shows behaviours and actions which can prevent cyberbullying. The two variable which show as important are smartphone usage and if the children believe they know more about Social Media than their parents. However, contained in the FPT is the *Complement* operator which makes interpreting the tree a little more difficult. This acts a logical reverse and flips the logic of the tree to that point. Therefore, an interpretation of this tree could be that not using a smartphone and not believing that you know more about Social Media than your parents will dramatically reduce the chance of you experiencing cyberbullying. Curbing or outright banning the use of smartphones in both primary and secondary schools has been a key government goal [28], and our analysis reinforces the need for this policy. As well as this, the need to increase parents and guardians knowledge of Social Media to curb cyberbullying has been a consistent theme throughout our analyses.

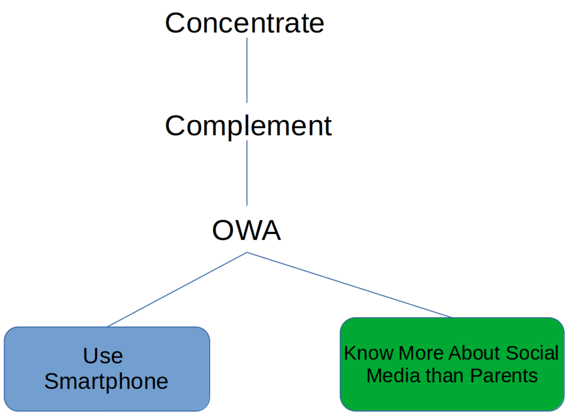

Figure 5: FPT used to Predict Absence of Cyberbullying. Green boxes show Attitudes while Blue boxes show behaviours. This tree shows that no Smartphone usage and not believing that they know more than their parents about social media reduce the risk of cyberbullying.

5. Conclusion

We collected and analysed the largest survey of cyberbullying in primary school children in Ireland. We successfully used this dataset to train an explainable AI classifier, called a Fuzzy Pattern Tree, to predict when cyberbullying will occur, achieving close to state-of-the-art accuracy. Fuzzy Pattern Trees can predict cyberbullying occurrence with an average of 84.3% accuracy, just behind the best method, Random Forest, with attains mean performance of 85.7%. Crucially however, Fuzzy Pattern Trees, as a white-box machine learning method, allow for their internal workings to be directly examined and their logic investigated. This ensures they can be safely deployed and also allows for specific actions to be undertaken based on the features they use for classification. The best fuzzy pattern tree model found suggests that more education is need for parents around the apps children use and their behaviour on social media. It also strongly suggests that reducing the use of smartphones among primary school students will reduce the risk of a child experiencing cyberbullying.

There are many future avenues for future research. The MCC scores of each of the classifiers were underwhelming and may point towards some slight over fitting that occurred during training. This requires investigation. How to best aggregate or combine multiple predictions from many different FPTs to improve results, while keeping expainability, also needs exploration. Finally, how to best use the model to automatically create personalised, actionable plans for students to help them best avoid cyberbullying necessitates study.

References

[1] G. D'Urso, J. Symonds, Risk factors for child and adolescent bullying and victimisation in ireland: A systematic literature review, Educational Review 75 (2023) 1464–1489.
[2] S. S. Ho, L. Chen, A. P. Ng, Comparing cyberbullying perpetration on social media between primary and secondary school students, Computers & Education 109 (2017) 74–84.
[3] G. W. Giumetti, R. M. Kowalski, Cyberbullying via social media and well-being, Current Opinion in Psychology 45 (2022) 101314.
[4] A. Cohen, A. Bendelow, T. Smith, C. Cicchetti, M. M. Davis, M. Heffernan, Parental attitudes on social media monitoring for youth: cross-sectional survey study, JMIR pediatrics and parenting 6 (2023) e46365.
[5] A. Murphy, M. Ali, D. M. Dias, J. Amaral, E. Naredo, C. Ryan, Grammar-based Fuzzy Pattern Trees for Classification Problems, in: Proceedings of the 12th International Joint Conference

on Computational Intelligence - ECTA„ INSTICC, SciTePress, 2020, pp. 71–80. doi:10.5220/0010111900710080.
[6] A. Murphy, G. Murphy, J. L. M. Amaral, D. M. Dias, E. Naredo, C. Ryan, Towards incorporating human knowledge in fuzzy pattern tree evolution, in: T. Hu, N. Lourenço, E. Medvet (Eds.), Genetic Programming - 24th European Conference, EuroGP 2021, Held as Part of EvoStar 2021, Virtual Event, April 7-9, 2021, Proceedings, volume 12691 of *Lecture Notes in Computer Science*, Springer, 2021, pp. 66–81. URL: https://doi.org/10.1007/978-3-030-72812-0_5. doi:10.1007/978-3-030-72812-0_5.
[7] M. Foody, M. Samara, J. O'Higgins Norman, Bullying and cyberbullying studies in the school-aged population on the island of ireland: A meta-analysis, British journal of educational psychology 87 (2017) 535–557.
[8] CyberSafeKids, Trends usage report 2022-23, https://www.cybersafekids.ie/wp-content/uploads/2023/11/CSK_Data-Trends-Report-2023-Web-Version-Final.pdf, 2023. Accessed : 2024.
[9] A. O'Connell, Coco's law: Two years on, Irish Criminal Law Journal 33 (2023) 26–37.
[10] A. Lonergan, A. Moriarty, F. McNicholas, T. Byrne, Cyberbullying and internet safety: a survey of child and adolescent mental health practitioners, Irish journal of psychological medicine 40 (2023) 43–50.
[11] B. A. Talpur, D. O'Sullivan, Cyberbullying severity detection: A machine learning approach, PloS one 15 (2020) e0240924.
[12] M. Fink, The eu artificial intelligence act and access to justice, EU Law live (2021) 1–4.
[13] C. Ryan, J. J. Collins, M. O. Neill, Grammatical Evolution: Evolving programs for an arbitrary language, in: European Conference on Genetic Programming, Springer, 1998, pp. 83–96.
[14] B. Majeed, S. Carvalho, D. M. Dias, A. Youssef, A. Murphy, C. Ryan, Performance Upgrade of Sequence Detector Evolution Using Grammatical Evolution and Lexicase Parent Selection Method, in: International Conference on Complex Computational Ecosystems, Springer, 2023, pp. 90–103.
[15] A. Murphy, T. Laurent, A. Ventresque, The case for grammatical evolution in test generation, in: Proceedings of the Genetic and Evolutionary Computation Conference Companion, 2022, pp. 1946–1947.
[16] M. Mahdinejad, A. Murphy, M. Tetteh, A. de Lima, P. Healy, C. Ryan, Grammar-guided evolution of the U-Net, in: International Conference on the Applications of Evolutionary Computation (Part of EvoStar), Springer, 2023, pp. 672–686.
[17] M. O'Neill, C. Ryan, M. Keijzer, M. Cattolico, Crossover in grammatical evolution, Genetic Programming and Evolvable Machines 4 (2003) 67–93. URL: https://doi.org/10.1023/A:1021877127167. doi:10.1023/A:1021877127167.
[18] N. Lourenço, F. B. Pereira, E. Costa, SGE: a structured representation for grammatical evolution, in: International Conference on Artificial Evolution (Evolution Artificielle), Springer, 2015, pp. 136–148.
[19] Z. Huang, T. D. Gedeon, M. Nikravesh, Pattern trees induction: A new machine learning method, Trans. Fuz Sys. 16 (2008) 958–970. doi:10.1109/TFUZZ.2008.924348.
[20] Y. Yi, T. Fober, E. Hüllermeier, Fuzzy operator trees for modeling rating functions, International Journal of Computational Intelligence and Applications 8 (2009) 413–428.
[21] A. Murphy, M. S. Ali, D. M. Dias, J. Amaral, E. Naredo, C. Ryan, Grammar-based fuzzy pattern trees for classification problems, in: Proceedings of the 12th International Joint Conference on Computational Intelligence - Volume 1: ECTA„ INSTICC, SciTePress, 2020, pp. 71–80. doi:10.5220/0010111900710080.
[22] A. Murphy, G. Murphy, D. M. Dias, J. Amaral, E. Naredo, C. Ryan, Human in the loop fuzzy pattern tree evolution, SN Computer Science 3 (2022) 163.
[23] A. Murphy, N. Lourenço, A. Ventresque, Initialisation in Structured Grammatical Evolution, in: Proceedings of the Companion Conference on Genetic and Evolutionary Computation, 2023, pp. 2022–2028.
[24] A. Murphy, A. Ventresque, C. Ryan, Fuzzy Pattern Trees with Pre-classification, in: International Conference on Complex Computational Ecosystems, Springer, 2023, pp. 104–117.

[25] L. Breiman, Random forests, Machine learning 45 (2001) 5–32.
[26] M. A. Hearst, S. T. Dumais, E. Osuna, J. Platt, B. Scholkopf, Support vector machines, IEEE Intelligent Systems and their applications 13 (1998) 18–28.
[27] T. Chen, C. Guestrin, Xgboost: A scalable tree boosting system, in: Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining, 2016, pp. 785–794.
[28] IrishTimes, Mobile phones set to be banned across all second-level schools under new Government plans, https://www.irishtimes.com/ireland/education/2024/08/21/minister-plans-mobile-phone-ban-across-second-level-schools/, 2024. Accessed : 2024.