UCD Michael Smurfit Graduate Business School

Assignment Topic	Classification and ML Ethics
Name	Anushka Jain
Student No.	22200246
Programme	MSc. Business Analytics
Module Title	Data Management and Mining - 2022/23 Spring
Module Code	MIS41270
Lecturer	Elayne Ruane
Submission Deadline	01st May 2023
Session	2022/23 Spring
Grade/Mark	

A SIGNED COPY OF THIS FORM MUST ACCOMPANY ALL SUBMISSIONS FOR ASSESSMENT. STUDENTS SHOULD KEEP A COPY OF ALL WORK SUBMITTED.

Procedures for Submission and Late Submission

Ensure that you have checked the school's procedures for the submission of assessments.

Note: There are penalties for the late submission of assessments. For further information please see the University's ***Policy on Late Submission of Coursework,*** (http://www.ucd.ie/registrar/)

Plagiarism: the unacknowledged inclusion of another person's writings or ideas or works, in any formally presented work (including essays, examinations, projects, laboratory reports or presentations). The penalties associated with plagiarism are designed to impose sanctions that reflect the seriousness of the University's commitment to academic integrity. Ensure that you have read the University's ***Briefing for Students on Academic Integrity and Plagiarism*** and the UCD ***Plagiarism Statement, Plagiarism Policy and Procedures***, (http://www.ucd.ie/registrar/)

Declaration of Authorship

I declare that all material in this assessment is my own work except where there is clear acknowledgement and appropriate reference to the work of others.

Name : Anushka Jain

Date : 01st May 2023

Table of Contents

Theory & Introduction

Data – a small word of just 4 letters but now the whole world is data-driven. It is the basis for AI and ML, where we use this data to understand society, make business choices, and attain growth. In ML, we have 3 types of learning: Supervised (predicts for defined labels), Unsupervised (finds hidden patterns), and Reinforcement Learning (Similar to human learning using trial and error) (Coursera, 2022).

Our main focus is Supervised Learning where human supervision is required to accurately label and train the algorithm (Ramakrishnan, 2022). The machine is trained with the aid of labeled input and output data to interpret the training data and predict precise results for fresh and untested data (Ramakrishnan, 2022). It can be either classifying data into a category (classification problem) or forecasting an outcome (regression algorithms) (Gong, 2022). For the assignment, we will be focusing on Classification.

Classification vs Regression (Gong, 2022): Classification models classify objects whereas regression models predict continuous outputs. The fine line between Classification and regression methods sometimes blurs. Classification can maybe just be regression along with a threshold. The number is true if it's above the threshold and false otherwise.

Classification:

In such problems, we try to fit or identify the category (already known to us) in which the new data will fit. It can be applied to both structured and unstructured data to precisely predict whether the data will fall into predetermined categories (Ramakrishnan, 2022). To assign it into various categories, we will have at least 2 or more than 2 categories. Classification uses probability score which it generates itself to assign it into a category (Ramakrishnan, 2022). Examples – whether an email is spam or not, image classification (whether the given image is a dog or a cat or another object), sentiment analysis (analyzing whether the input text is positive, negative, or neutral), disease diagnosis, etc.

There are various Classification Predictive Modelling Algorithms (Ramakrishnan, 2022):

1. Binary classification: logistic regression, decision trees, simple bayes & support vector machine
2. Multi-Class classification: k nearest neighbors, choice trees
3. Multi-Label classification: multi-label random forest
4. Imbalanced classification

Task – 1

1.1 MELD: A multimodal multi-party dataset for emotion recognition in conversations

For my first research paper review, I chose the second paper from the given list i.e., "MELD: A multimodal multi-party dataset for emotion recognition in conversations" by Poria, Soujanya, et al. from 2018. After going through (Keshav, 2007) suggested by Elayne, as to how to read the various research papers, I applied the same 3-pass method here. Details of the first pass for this paper is as follows:

Category: This work introduces a new multimodal dataset, MELD, for emotion recognition in spoken language research. Several baseline models' results on the dataset are included as benchmarks as well.

Context: This work expands upon prior studies in the domain of conversational emotion recognition. Several theoretical frameworks have been employed by the authors in their investigation of the emotion recognition in conversations problem. These frameworks include the appraisal theory of emotions and multimodal affective computing.

Correctness: The paper's underlying assumptions make sense to me. The authors have amassed a huge and varied dataset, and they used a number of human annotators to guarantee accurate emotion identification. The usefulness of utilizing several modalities in emotion recognition is further validated by the benchmark results obtained on the dataset.

Contributions: This study's primary contributions are the creation of a new multimodal dataset for conversational emotion recognition and the benchmarking of different baseline models on that dataset. This work emphasizes the necessity of including several modalities while conducting emotion recognition research, and it offers important guidance for researchers in the field.

Clarity: The paper is well-written, with clear and concise language. The authors have provided detailed descriptions of the dataset collection process and annotation methodology, making it easy to understand and replicate their work. The benchmark results are also presented in a clear and organized manner, making it easy to compare the performance of different models. Overall, the paper is well-organized and easy to follow.

Summary

The surprising element in the paper is the inclusion of complex emotions, such as sarcasm and excitement, in the emotion categories. The paper focuses on a multimodal approach to emotion recognition, combining information from text, audio, and visual modalities. However, alternative approaches, such as using deep learning models, could also be explored. I found the technical details of the benchmark results section difficult to understand, particularly because I was not familiar with the specific models used. The clear and detailed descriptions of the dataset collection process and annotation methodology aided in understanding the paper. Using the method of interparty dependency modeling proposed by Majumder et al. (2019), they created a robust baseline capable of emotion recognition in multi-party interactions.

Dataset reannotation was carried out for increased accuracy of the dataset. This changed the annotation based on facial expressions. These examples highlight the significance of context and multimodality in recognizing emotions during speech. Non-neutral aka disagreement in annotating the utterances show the reality of the dataset. IEMOCAP and SEMAINE contain dyadic (between 2 people only) conversations, while MELD is multi-party. Further, since everything was relatively new for me, I referred to (elvis, 2019) and (www.mathworks.com, n.d.) for better understanding of Dialogue RNN and convolutional neural network (CNN or ConvNet) respectively.

The paper is well written, and it has increased my knowledge and understanding of sentiment and emotion analysis. I realized how important context is and used RNNs for its modelling, as the emotional factors depend on both what was said before and how it made the person feel. Multiple studies have benefited from publicly available datasets for multimodal emotion recognition in conversations, such as IEMOCAP and SEMAINE. However, these datasets had limitations, such as a lack of multi-party conversations and a relatively low number of total utterances, thus creation of MELD was important. When compared to sentiment classification, emotion classification performs poorly. As emotion classification is concerned with more nuanced classes, this finding makes intuitive sense. When it comes to real-time personal assistants like Siri and Google Assistant, where users can use speech, text, and facial expressions to communicate, multimodal dialogue systems can be extremely useful.

Limitations

The main limitation of the paper is that the dataset is primarily focused on English conversations and may not generalize well to other languages or cultures. Additionally, the dataset is relatively small compared to some other datasets used in emotion recognition research. Collecting a similar dataset with a varied variety of emotions, individuals, and circumstances can be difficult, especially in a different cultural or language context. Replication requires high-quality emotional state annotations. Finding expert annotators who can appropriately classify emotions in discussions may be difficult. Integrating text, audio, and visual cues involves careful consideration of each modality's features and methodologies, as well as its weighting. Selecting effective models for emotion recognition, especially multimodal models, needs expertise in machine learning, natural language processing, and model architectural strengths and weaknesses.

Overall Impressions

The paper could be improved by including more detailed explanations of the theoretical bases used to analyze the problem of emotion recognition in conversations. Additionally, a larger and more diverse dataset could be collected in the future to address some of the limitations of the current dataset. The paper was published on arXiv, a repository of scientific research papers, and does not appear to have been formally published in a peer-reviewed journal. However, on further research, I found that the authors are well-known researchers in the field of natural language processing and machine learning. However, none of these factors influenced my decision to provide a summary of this particular research paper. Although it used a dataset from one of my favorite TV shows, which may have influenced my decision to choose this particular paper, the concept of ERC piqued my interest. This is also because our main goal with ML is to make the computer as human as possible, therefore the most important attribute that people have are emotions, which, if understood and recreated by a machine, may achieve miracles. Finally, sentiment analysis has been discussed frequently during lectures, which intrigued my interest enough to select this paper.

1.2 Application of Random Forest Algorithm on Feature Subset Selection and Classification and Regression

For my second paper, I chose (Jaiswal and Samikannu, 2017)'s Random Forest Algorithm. Because the first paper did not go into detail about any of the classification algorithms, I wanted to choose my second paper so that I could reflect on my learnings. I also applied the three-pass strategy (Keshav, 2007), the details of which are described below.

Category: This study explains the use of the Random Forest algorithm for feature subset selection and classification & regression problems.

Context: The work is connected to studies that investigate using machine learning methods for feature selection and classification/regression. Analysis of the issue is grounded in theory and makes use of tools like decision trees, ensemble approaches, and feature selection strategies.

Correctness: It looks like the paper's assumptions are correct. The RF algorithm's accuracy, sensitivity, specificity, and AUC-ROC are all measured and compared to those of other machine learning methods across many datasets in this study. The outcomes demonstrate that the RF algorithm excels in feature subset selection and classification/regression tasks, where other algorithms fall short.

Contributions: The primary contributions of this study are an analysis of the performance of the RF method and comparisons to other machine learning algorithms for use in feature subset selection and classification/regression tasks. The research demonstrates that the RF algorithm is a versatile and powerful method for dealing with huge datasets that contain complex features.

Clarity: Excellent writing skills are on display throughout this paper's introduction, methods, findings, and conclusions. The study's methodology is described, and the report gives a thorough account of the results, complete with tables and figures. The findings in the paper lend credence to the findings.

Summary

The paper explores the application of Random Forest (RF) algorithm on feature subset selection and classification/regression tasks. The study aims to evaluate the effectiveness of RF in feature selection and classification/regression tasks, and to compare its performance with other machine learning algorithms such as Support Vector Machines (SVM), Artificial Neural Networks (ANN), and k-Nearest Neighbors (k-NN).

The study employs several datasets, including the Chronic Kidney Disease, the Air Quality dataset, to evaluate the performance of the RF algorithm in feature subset selection and classification/regression tasks. The RF algorithm is compared with other machine learning algorithms in terms of accuracy, sensitivity, specificity, and area under the receiver operating characteristic curve (AUC-ROC).

Despite its apparent simplicity, the Random Forest approach and outperforms other machine learning algorithms on feature subset selection and classification/regression tasks, which is one of the paper's most striking findings. A feature that will allow me to divide the observations into groups that are as distinct from one another as possible, while maintaining as much similarity as possible within each resulting subgroup. The performance of a committee made up of many relatively uncorrelated models (trees) will always exceed that of any single model in the committee. Even though the paper was detail oriented with basics, I still referred to a lot of online articles for further understanding. Like, for Gini index, which is the measure of variance I referred to (Tat, 2017). More mis-classification occurs when the variance is larger. A lower Gini Index value indicates more accurate classification (Tat, 2017).

Limitations

The study has limitations in terms of the datasets used for evaluation, as they are limited to a specific domain. The study could have included more diverse datasets from different domains to improve the generalizability of the results. The paper does not provide detailed information on the specific parameters used in the Random Forest algorithm, which may make it difficult to reproduce the exact results presented in the study. The paper briefly mentions other machine learning algorithms used for feature subset selection and classification/regression tasks, such as decision trees, support vector machines, and logistic regression. However, the study does not provide a detailed comparison of the performance of these algorithms with the Random Forest algorithm. Overall, replicating the work presented in the paper may require a significant amount of time, effort, and resources, and may be challenging for researchers without sufficient expertise or resources.

Overall Impressions

I am mostly intrigued by the intuitiveness of the random forest approach. The fact that low correlation between models is the key, as the chances of making correct predictions increase with the number of uncorrelated trees in our model. To ensure that they aren't too correlated, they use Bagging (sample with replacement) and feature randomness. They could have improved by keeping the font size and style same across the entire paper, the formatting could have been improved. Overall, the paper is well-written and easy to follow, and there were no specific parts that were difficult to understand. The paper's clear organization and presentation of results aided me in my reading. The methodology and results were presented in a logical and concise manner, making it easy to follow. The paper could be improved by including more diverse datasets from different domains to improve the generalizability of the results. Additionally, a more detailed comparison of the performance of the Random Forest algorithm with other machine learning algorithms could be included. The paper's conclusions are well-supported by the results presented in the study. The study shows that the Random Forest algorithm is an effective tool for feature subset selection and classification/regression tasks and can be applied to various domains. The paper was published by IEEE Xplore. The publisher did not have a significant impact on my understanding of the research, as the paper's content was well-presented and easy to follow.

Task – 2.1

1. Quantifying the Carbon Emissions of Machine Learning

The research paper "Quantifying the carbon emissions of machine learning" by Lacoste et al. aims to quantify the carbon emissions associated with training deep learning models. The authors note that there has been increasing concern over the energy consumption and carbon footprint of machine learning, particularly in light of the growing popularity and use of deep learning models. Continuing the literary reading methodology (Keshav, 2007) developed, the details of first-pass component are as follows:

Category: in my opinion, the paper can be categorized as an analysis paper.

Context: This study is connected to the urgent need to address the environmental impact of technology. The authors have analyzed the carbon emissions of machine learning algorithms by drawing on past studies on the energy consumption and carbon emissions of technology.

Correctness: The assumptions made by the authors appear to be valid. The authors have used a methodical approach to quantify the carbon emissions of machine learning, taking into account various factors such as hardware, training time, cloud provider and region of computation.

Contributions: The enormous carbon footprint of machine learning models is demonstrated, and a systematic way is presented to quantify these emissions. The report also stresses the importance of environmentally conscious design and sustainable machine learning methods.

Clarity: The work is nicely written and easy to understand, both in terms of its findings and its methods. The authors have included graphical and tabular representations of the findings to enhance the reader's comprehension.

Summary

To address this issue, the authors developed a framework for estimating the carbon emissions associated with training deep learning models. The framework (Lacoste et al., 2019) takes into account factors such as the energy consumption of the hardware used for training, the carbon intensity of the electricity used, and the number of training iterations required.

The authors suggest that machine learning practitioners can reduce the carbon footprint of their models by using more energy-efficient hardware, optimizing hyperparameters, reducing model complexity, and using more energy-efficient algorithms. The availability of actionable items for an individual or the organization to take post analysis was one of the interesting takes. Another interesting piece of information was the fact that GCP and Microsoft Azure are 100% carbon neutral while AWS is not. Somehow, this new information will anchor me towards choosing GCP and Azure over AWS. The authors then used this framework to estimate the carbon emissions of training several popular deep learning models, including ResNet-50 and BERT. Their results

showed that the carbon emissions of training these models can range from a few hundred pounds of CO_2 for smaller models to several tons of CO_2 for larger models.
The paper concludes by highlighting the importance of considering the carbon emissions of machine learning and encouraging researchers and practitioners to be more mindful of the environmental impact of their work. The authors also suggest several strategies for reducing the carbon footprint of machine learning, such as using more energy-efficient hardware, optimizing model architecture to reduce the number of training iterations required, and using renewable energy sources. Overall, the paper provides an important contribution to the growing discussion around the environmental impact of technology and the need for more sustainable practices in the field of machine learning.

Limitations

The main limitation of this paper is that it focuses only on the energy consumption of training and deploying machine learning models and does not consider the potential benefits that these models can bring, such as reducing energy consumption in other areas. Additionally, the authors note that their study does not consider the impact of advancements in hardware or algorithmic efficiency on reducing carbon emissions. The calculations for carbon emissions require significant computational resources and are time-consuming. Therefore, researchers might need to have access to high-performance computing facilities and thus might face issue in reproducing the work. Finally, the study's carbon footprint analysis relies on carbon intensity data, which might not be available for all regions or might not be up to date.

Overall Impressions

As part of Task 1 and in the previous assignment, I went through a lot of algorithms, trying to understand their methodology and their pros and cons. But I never thought of the environmental impact it could have. Thus, the most surprising element of this paper is the extent to which machine learning models contribute to carbon emissions, as well as the potential for reducing emissions by optimizing these models.

Some of the technical details in the paper, such as the metrics used to quantify carbon emissions and the specific hardware and software configurations used, were difficult to understand for those without a background in machine learning or energy consumption analysis. The paper's clear organization, with proper explanation and references to illustrate key points aided in understanding. The paper's discussion and conclusions were well-supported by the research presented, and the authors provide useful recommendations for reducing the carbon footprint of machine learning models. Another interesting take was the inclusion of near real-time public data. The fact that the authors are keen on updating the database for continued research is like the green flag of a research paper. Further, they are transparent and also talk about the large existence of error because of lack of public information of the real usage of energy and emissions. The paper was published by researchers from Mila, Element AI, and University of Sherbrooke. Knowing the backgrounds of the authors and their affiliation with reputable organizations provided confidence in the credibility of their findings.

2. Machine Bias

It is without a doubt one of the most informative and thought-provoking articles that I have ever read. After reading ProPublica article (Angwin et al., 2016) from 2016, I feel curious as to how much of my environment is affected by AI and biases in Machines and Algorithms. I have continued the strategy of 3-pass method, even though it is an article and not a research paper. Details of which are listed below:

Category: Journalistic in nature, the paper summarizes the results of a study.

Context: Several other studies and sources are cited to set the stage for the problem of machine bias in risk assessments for criminal sentencing.

Correctness: The article details the findings of a study by ProPublica that has been widely discussed in the press and the academy. However, there may be varying opinions on the validity of the study's findings because the problem of machine bias is a contentious and ongoing topic.

Contributions: It brings attention to the problem of machine prejudice in the criminal justice system and emphasizes the hazards of depending on automated technologies without adequate testing and supervision.

Clarity: This post is really well written, and the author avoids jargon in favor of simple explanations.

<u>Summary</u>

The article "Machine Bias" by ProPublica journalists delves at the use of machine learning algorithms to forecast the likelihood of recidivism among convicted criminals. The essay focuses on how these algorithms can perpetuate racial bias in the criminal justice system, especially against African Americans.

The journalists investigated Northpointe, the makers of the widely used risk assessment program COMPAS, to learn more about the algorithms utilized by the corporation. Compared to white defendants, black defendants were more likely to be wrongly flagged as being at a higher risk of recidivism by COMPAS, the inquiry found. Several instances where these predictions were proved wrong and resulted in harsher sentences for black defendants are provided in the article.

The authors also conducted interviews with specialists who shed light on the technological underpinnings of these biases, such as the utilization of historical data that reflects systemic discrimination against minorities. They worry that the use of such algorithms will only serve to reinforce the criminal justice system's preexisting racial biases. As Brennan (creator of COMPAS) mentioned (Angwin et al., 2016):

“If those are omitted from your risk assessment, accuracy goes down,”

Those: items that can be correlated with race — such as poverty, joblessness and social marginalization

Another aspect of the article that caught my attention was Zilly's understanding that the scores did not reflect the positive changes he was making in his life, such as his conversion to Christianity, his struggle to quit drugs, and his efforts to be more present for his son. At the end of the piece, the authors make a plea for more transparency in the implementation of such algorithms and the creation of more fair and precise tools. It emphasizes the significance of scrutinizing the data used to train machine learning algorithms and being cognizant of the possibility of biases present. Ultimately, the article underlines the need for greater scrutiny and responsibility in the development and deployment of machine learning algorithms used in the criminal justice system and poses significant questions about their use.

Limitations

The article is an investigative report. Replicating the investigation described in the article would likely require access to extensive criminal justice data, as well as expertise in data analysis and investigative journalism. Given the delicate nature of the topic at hand, there may also be legal and ethical hurdles to overcome in order to gain access to the relevant information and carry out the inquiry.

Overall Impressions

Using risk assessment algorithms in criminal sentencing has been shown to be discriminatory against some groups, including Black defendants, as this article explains. The piece also features stories of those whose lives were altered by these prejudices. When Judge John Hurley of Broward County first started out, he relied on COMPAS (Correctional Offender Management Profiling for Alternative Sanctions) rather than relying on his own judgment. Because of this predisposition, I think all he needed was to be anchored in place. However, Brennon (the creator of COMPAS) stated that he did not approve of the idea of using COMPAS as the only evidence in making decisions. His goal was criminal prevention, not moral judgment. It scares me to think about how skewed the decisions have been as AI is becoming ubiquitous and our reliance on it with blind confidence that it would be right is increasing with many folds. Possible drawbacks of utilizing machine learning algorithms in the criminal justice system are discussed in depth, providing food for thought. It stresses the need for developers and implementers of such algorithms to think about bias and fairness. I had doubts about the reliability of the system because its creators doubled as the Quality Assurance (QnA) Team. Finally, the judges were motivated to use Northpointe's software despite the possibility of bias because it was simple to use, produce, and provide effective charts and graphs for judicial review.

The article is written in accessible language and presents the information in a clear manner. The use of real-life examples and personal stories helped to illustrate the impact of machine bias in a tangible way. This investigative piece was published by ProPublica, a nonprofit news organization. I had more faith in the article's veracity because the publication was published by a reputable outlet known for its investigative reporting.

Task – 2.2

Timnit Gebru is a well-respected computer scientist and AI researcher (Wikipedia, 2023) who has made significant contributions to the fields of AI ethics and society, algorithmic bias, and tech diversity and inclusion, with contributions to the fields of computer vision and natural language processing (NLP).

In the realm of artificial intelligence (AI), Gebru has made major contributions, especially in the areas of addressing biases and assuring ethical practices in AI algorithms and systems. Her studies on AI's repercussions on society have shed light on the importance of including underrepresented groups in the field. As an ardent advocate of diversity and inclusion in tech, Gebru also helped found the organization, Black in AI, which works to elevate the profile of Black researchers.

Gebru's work on the effects of algorithmic prejudice on marginalized communities is well-known. Both "Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification" and "Datasheets for Datasets" are two of her seminal works on the subject. In their article "Gender Shades," co-authors Gebru and Joy Buolamwini found that facial recognition algorithms performed significantly worse on women with darker skin tones than on men (Buolamwini and Gebru, 2018). The media's attention to this study ultimately prompted demands for greater openness and responsibility in the creation of AI systems.

Unfortunately, I was unable to obtain access to the article "Gender Shades" that I had hoped to summarize for Task 2.1. I intend to read all of Gebru's works in depth after learning about her path in ML Ethics. Like, the proposed datasheet in the paper “Datasheets for Datasets” (Gebru et al., 2018) includes information such as dataset composition, collection process, recommended uses, and ethical considerations.

Task – 3

ML Ethics Understanding

Machine Learning Ethics or Artificial intelligence ethics is a code of conduct that provides direction for the creation and use of AI. Cognitive biases, such as recency and confirmation bias, are built into the human brain and manifest themselves in our actions and, ultimately, in the data we collect (www.ibm.com, n.d.).

In our subject, Decision and Behavioural Analytics, I learned about a wide range of heuristics and biases, such as representativeness, availability, endowment effect, anchoring, design decisions, and many others. Humans have introduced biases into the data, which the computers then normalize through their learnings (supervised, unsupervised, and reinforced), despite their best intentions and efforts to make machines as human-like as feasible. It's possible that the "correct" data being put into the algorithms is actually biased and twisted. Imagine a hiring manager who is convinced that only those with degrees from the best universities will be the most effective workers; upon learning that the candidate in question holds one of these degrees, the manager may look for evidence to back up his preconceived notions (confirmation bias) (Jones, 2019). While such data would make for an accurate dataset, it could cause havoc if used to train machines to make similar predictions.

As a result, I think it's crucial to factor in ethics, and not just in the areas of privacy, security, consent, and GDPR.

Influence of ML Ethics on MELD - ERC

ML Ethics is concerned with ensuring that these technologies are developed and used in a responsible and ethical manner. ML models like MELD in ERC (Multimodal Emotion Recognition Challenge in Emotion Recognition in Conversations) must adhere to ethical guidelines to ensure that they do not perpetuate bias, discrimination, or harm to individuals or groups.

MELD in ERC is a multimodal emotion recognition dataset that contains audio, video, and textual data for emotion recognition in conversations. Ethical concerns in MELD in ERC are related to issues of privacy, consent, and bias. Privacy concerns arise when sensitive personal data is collected and shared without the knowledge or consent of individuals. In MELD in ERC, FRIENDS' database has been used with proper explicit consent. Thus, there are no ethical concerns around privacy and the right to control one's personal data.

However, bias in the dataset and the ML models trained on it presents a potential ethical problem for MELD in ERC. Lack of diversity in the dataset or bias in the algorithms used to evaluate it can lead to inaccurate results. Furthermore, there is no data on the potential biases that permeated the different annotators. Despite the incorporation of re-annotation, we have no idea if cultural biases were at play or if the annotators represented a wide range of backgrounds.

Joey on FRIENDS is an Italian American, while Chandler Bing is of Scottish descent. The authors, the characters they created, and the people who annotated their work may have all come from the same cultural background, but that doesn't mean they're all thinking and feeling the same way. So, how can we be sure that the understanding has been proper. This could lead to biased or incorrect conclusions, which would only serve to reinforce preexisting prejudices.

Researchers need to think about the ethical consequences of their work and take precautions to ensure ethical standards in ML models like MELD in ERC. This involves being forthright about the development and operation of the models, receiving informed consent from participants, ensuring diversity in the dataset, reviewing algorithms for bias, and more.

REFERENCES

1. Angwin, J., Larson, J., Mattu, S. and Kirchner, L. (2016). Machine Bias. [online] ProPublica. Available at: https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing.
2. aigents.co. (n.d.). Introduction to Convolutional Neural Networks CNNs. [online] Available at: https://aigents.co/data-science-blog/publication/introduction-to-convolutional-neural-networks-cnns.
3. Coursera (2022). *3 Types of Machine Learning You Should Know*. [online] Coursera. Available at: https://www.coursera.org/articles/types-of-machine-learning.
4. elvis (2019). DialogueRNN: Emotion Classification in Conversation. [online] DAIR.AI. Available at: https://medium.com/dair-ai/dialoguernn-emotion-classification-in-conversation-1e389d035aff#:~:text=DialogueRNN%20is%20a%20method%20that.
5. Gebru, T., Morgenstern, J., Vecchione, B., Vaughan, J.W., Wallach, H., Daumeé III, Hal and Crawford, K. (2018). Datasheets for Datasets. [online] arXiv.org. Available at: https://arxiv.org/abs/1803.09010.
6. Gong, D. (2022). Top 6 Machine Learning Algorithms for Classification. [online] Medium. Available at: https://towardsdatascience.com/top-machine-learning-algorithms-for-classification-2197870ff501.
7. J. Buolamwini and T. Gebru. 2018. "**Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification.**" In Proceedings of the 2018 Conference on Fairness, Accountability and Transparency, Vol. 81. PMLR, New York, NY, USA, 77–91. http://proceedings.mlr.press
8. Jaiswal, J.K. and Samikannu, R. (2017). Application of Random Forest Algorithm on Feature Subset Selection and Classification and Regression. [online] IEEE Xplore. doi:https://doi.org/10.1109/WCCCT.2016.25.
9. Jones, P. (2019). Unconscious bias – avoidable or inevitable? [online] ERC: European Research Council. Available at: https://erc.europa.eu/news-events/magazine/unconscious-bias-%E2%80%93-avoidable-or-inevitable.
10. Keshav, S. (2007). How to read a paper. ACM SIGCOMM Computer Communication Review, [online] 37(3), p.83. doi:https://doi.org/10.1145/1273445.1273458.
11. Lacoste, A., Luccioni, A., Schmidt, V. and Dandres, T. (2019). Quantifying the Carbon Emissions of Machine Learning. arXiv:1910.09700 [cs]. [online] Available at: https://arxiv.org/abs/1910.09700.

12. Poria, S., Hazarika, D., Majumder, N., Naik, G., Cambria, E. and Mihalcea, R. (n.d.). MELD: A Multimodal Multi-Party Dataset for Emotion Recognition in Conversations. [online] Available at: https://arxiv.org/pdf/1810.02508.
13. Ramakrishnan, M. (2022). What is Classification in Machine Learning and Why is it Important? [online] Emeritus Online Courses. Available at: https://emeritus.org/blog/artificial-intelligence-and-machine-learning-classification-in-machine-learning/#:~:text=In%20machine%20learning%2C%20classification%20is.
14. Tat, M.J. (2017). Seeing the random forest from the decision trees: An explanation of Random Forest. [online] Medium. Available at: https://towardsdatascience.com/seeing-the-random-forest-from-the-decision-trees-an-intuitive-explanation-of-random-forest-beaa2d6a0d80#:~:text=The%20Gini%20Index%20is%20a.
15. Wikipedia. (2023). Timnit Gebru. [online] Available at: https://en.wikipedia.org/wiki/Timnit_Gebru#.
16. www.ibm.com. (n.d.). AI Ethics | IBM. [online] Available at: https://www.ibm.com/topics/ai-ethics#:~:text=AI%20ethics%20is%20a%20set.
17. www.mathworks.com. (n.d.). What is a Convolutional Neural Network? [online] Available at: https://www.mathworks.com/discovery/convolutional-neural-network-matlab.html#:~:text=A%20convolutional%20neural%20network%20(CNN.
18. Yiu, T. (2019). Understanding Random Forest. [online] Medium. Available at: https://towardsdatascience.com/understanding-random-forest-58381e0602d2.