REPORT ON SUMMER INTERNSHIP

Fair classification via Crowdsourcing

Under the Technical Institute

"INDIAN INSTITUTE OF TECHNOLOGY, ROPAR"

CONTINUING EDUCATION & OUTREACH ACTIVITIES

Rupnagar, Punjab- 140001, India

Phone: 01881- 231120/ email- office-ceora-1@iitrpr.ac.in

Submitted by

ROHAN GUPTA (18103047)

DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING
INSTITUTE OF TECHNOLOGY, GURU GHASIDAS
VISHWAVIDYALAYA, BILASPUR -495009

TABLE OF CONTENTS

Page No.

COMPANY PROFILE

Indian Institute of Technology Ropar
भारतीय प्रौद्योगिकी संस्थान रोपड़

Indian Institute of Technology Ropar (**IIT Ropar**) is a technical university located in Rupnagar, Punjab, India. It is one of the eight newer Indian Institutes of Technology (IITs) established by the Ministry of Human Resource Development (MHRD), Government of India under The Institutes of Technology (Amendment) Act, 2011 to expand the reach and enhance the quality of technical education in the country.

HISTORY

The Indian Institutes of Technology (IITs) are autonomous public technical universities located across India. It is under the jurisdiction of Ministry of Education, Government of India. They are governed by the Institutes of Technology Act, 1961, which has declared them as Institutes of National Importance and lays down their powers, duties, and framework for governance. The Institutes of Technology Act, 1961 lists twenty-three institutes. Each IIT is autonomous, linked to the others through a common council (IIT Council), which oversees their administration. The Minister of Education is the ex officio Chairperson of the IIT Council. As of 2021, the total number of seats for undergraduate programs in all IITs is 16,232.

The history of the IIT system nearly dates back to 1946 when Sir Jogendra Singh of the Viceroy's Executive Council set up a committee whose task was to consider the creation of *Higher Technical Institutions* for post-war industrial development in India. The 22-member committee, headed by Nalini Ranjan Sarkar, recommended the establishment of these institutions in various parts of India, along the lines of the Massachusetts Institute of Technology, with affiliated secondary institutions.[17]

The first Indian Institute of Technology was founded in May 1950 at the site of the Hijli Detention Camp in Kharagpur, West Bengal.[18] The name "Indian Institute of Technology" was adopted before the formal inauguration of the institute on 18 August 1951 by Maulana Abul Kalam Azad.[19] On 15 September 1956, the Parliament of India passed the *Indian Institute of Technology (Kharagpur) Act*, declaring it as an Institute of National Importance. Jawaharlal Nehru, first Prime Minister of India, in the first convocation address of IIT Kharagpur in 1956 said:[20]

Here in the place of that Hijli Detention Camp stands the fine monument of India, representing India's urges, India's future in the making. This picture seems to me symbolical of the changes that are coming to India.

On the recommendations of the Sarkar Committee, four campuses were established at Bombay (1958), Madras (1959), Kanpur (1959), and Delhi (1961). The location of these campuses was chosen to be scattered throughout India to prevent regional imbalance.[21] The *Indian Institutes of Technology Act* was amended to reflect the addition of new IITs.[1] Student agitations in the state of Assam made Prime Minister Rajiv Gandhi promise the creation of a new IIT in Assam. This led to the establishment of a sixth institution at Guwahati under the Assam Accord in 1994. In 2001, the University of Roorkee was converted into IIT Roorkee.[9] Over the past few years, there have been a number of developments toward establishing new IITs. On 1 October 2003, Prime Minister Atal Bihari Vajpayee announced plans to create more IITs "by upgrading existing academic institutions that have the necessary promise and potential".[22] Subsequent developments led to the formation of the S K Joshi Committee, in November 2003, to guide the selection of the five institutions which would be converted into IITs. Based on the initial recommendations of the Sarkar Committee, it was decided that new IITs should be spread throughout the country. When the government expressed its willingness to correct this regional imbalance, 16 states demanded IITs. Since the S K Joshi Committee prescribed strict guidelines for institutions aspiring to be IITs,[23] only seven colleges were selected for final consideration.[24] Plans are also reported to open IITs outside India, although there has not been much progress in this regard.[25] Eventually in the 11th Five year plan, eight states were identified for establishment of new IITs.

In 2008 to 2009, eight new IITs were set up in Gandhinagar, Jodhpur, Hyderabad, Indore, Patna, Bhubaneswar, Ropar, and Mandi. Following same selection process since 1972, in 2012 the Institute of Technology, Banaras Hindu University was made a member of the IITs and renamed as IIT (BHU) Varanasi.[10]

In 2015 to 2016, six new IITs in Tirupati, Palakkad, Dharwad, Bhilai, Goa, and Jammu, approved through a 2016 bill amendment, were founded, along with the conversion of Indian School of Mines Dhanbad into IIT (ISM) Dhanbad.[12]

The entire allocation by the central government for 2017-18 budget for all Indian Institutes of Technology (IITs) was slightly over ₹70 billion (US$930 million). However, the aggregate money spent by Indian students for tertiary education in the United States was about six times more than what the central government spends on all IITs.[26]

AWARDS AND RECOGNITION

IIT Ropar has ranked on 205th position in the QS Asia Ranking 2020.[26] The *Times Higher Education World University Rankings* ranked it 301–350 globally in the 2020 ranking.[27] The National Institutional Ranking Framework (NIRF) ranked it 25th among engineering colleges in 2020 and 39th in Overall ranking in NIRF 2020.

CERTIFICATE

INDIAN INSTITUTE OF TECHNOLOGY ROPAR

CONTINUING EDUCATION & OUTREACH ACTIVITIES

Internship Certificate

Ref. No. IITRPR/CEOA/SI/2021/29 **Dated: 17/08/2021**

This is to certify that **Mr/Ms. Rohan Gupta** a student of B.Tech from **Institue-Institute of Technology, Guru Ghasidas Vishwavidyalaya, Koni, Bilaspur** has done his/her Internship, titled **"Fair classification via crowdsourcing"** from **17/05/2021 to 17/07/2021** under the supervision of **Dr. Shweta Jain**, Assistant Professor/Associate Professor/Professor, Department of **Computer Science and Engineering** at this Institute.

18/8/2021

Dr. Shweta Jain	Dr. Nitin Auluck	Dr. Apurva Mudgal
Supervisor	**Head of Department**	**Associate Dean (CEOA)**

Rupnagar, Punjab-140001, India
Phone: 01881-231120 /e-mail: office-ceora-1@iitrpr.ac.in

DECLARATION

I hereby declare that all the work presented in this report in the partial fulfilment of the requirement for the award of the degree of Bachelor of Technology in Computer Science & Engineering, Institute of Technology, Guru Ghasidas Vishwavidyalaya, Central University, Bilaspur, Chhattisgarh, is an authentic record of the work done during the summer internship under IIT Ropar.

Rohan Gupta(18103047)

ACKNOWLEDGEMENT

The satisfaction that accompanies the successful completion of any task would be incomplete without the mention of people whose ceaseless cooperation made it possible, whose constant guidance and encouragement crown all efforts with success.

I would like to express my gratitude and appreciation to the **IIT Ropar, Department of Computer Science and Dr. Shweta Jain(Supervisor)** for being our guide and all those who gave us the opportunity to complete this training.

We are also deeply thankful to all sources from where we have cited information. We don’t know all of the names of people behind them, but we want to acknowledge their help during our internship.

ABSTRACT

This internship has allowed us to gain insights on study of research paper. The platform that we have worked upon here is Jupyter Notebook.

I have learnt what are the uses of research paper in our daily life problem and tools like Python, Machine Learning, Jupyter Notebook to create a fair dataset.

This project done by me was to complete my internship task with the implementation of the research paper where I developed a machine learning model on Adults Dataset based upon Random Forest Classifer and Analysis the dataset with fair classification as a crowdworker.The Final task was to analysis the sensitive attribute with ground truth labels. This project is a mid level project and this research paper has a wide scope of future enhancement.

Research Paper

Crowdsourcing with Fairness, Diversity and Budget Constraints

Naman Goel
naman.goel@epfl.ch
Artificial Intelligence Lab, EPFL
Lausanne, Switzerland

Boi Faltings
boi.faltings@epfl.ch
Artificial Intelligence Lab, EPFL
Lausanne, Switzerland

ABSTRACT

Recent studies have shown that the labels collected from crowdworkers can be discriminatory with respect to sensitive attributes such as gender and race. This raises questions about the suitability of using crowdsourced data for further use, such as for training machine learning algorithms. In this work, we address the problem of fair and diverse data collection from a crowd under budget constraints. We propose a novel algorithm which maximizes the expected accuracy of the collected data, while ensuring that the errors satisfy desired notions of fairness. We provide guarantees on the performance of our algorithm and show that the algorithm performs well in practice through experiments on a real dataset.

CCS CONCEPTS

• **Information systems** → **Crowdsourcing**; *Social networks*; Incentive schemes.

KEYWORDS

Crowdsourcing, Data Quality, Bias, Fairness

ACM Reference Format:

Naman Goel and Boi Faltings. 2019. Crowdsourcing with Fairness, Diversity and Budget Constraints. In *AAAI/ACM Conference on AI, Ethics, and Society (AIES '19), January 27–28, 2019, Honolulu, HI, USA.* ACM, New York, NY, USA, 8 pages. https://doi.org/10.1145/3306618.3314282

1 INTRODUCTION

Algorithmic decision-making is gaining popularity in many diverse application areas of social importance. Examples include criminal recidivism prediction, stop-and-frisk programs, university admissions, bank loan decisions, screening job candidates, fake news control, information filtering(personalization) and search engine rankings etc. Recently, questions were raised about the fairness of these algorithms. An investigation [26] found COMPAS (a popular software used by courts to predict criminal recidivism risk) racially discriminatory. Other software systems have also been found to be biased against people of different races, genders and political views [3, 18, 20, 25]. This has led to a widespread and legitimate concern about the potential negative influence of such systems on the society [2, 29]. One of the main reasons of algorithmic bias is the bias in the training datasets. In order to achieve algorithmic fairness, the issue of **data fairness** needs to be addressed first. In many interesting cases, data is directly or indirectly influenced by some kind of human feedback. The influence is obvious and direct if human assigned labels are used as a proxy for ground truth labels. However, human feedback can also indirectly influence the so-called "ground-truth" datasets (when the labels are not human assigned but observed in reality). This is because the ground truth labels can only be collected for a finite number of data points and the selection of data points is often influenced by humans. For example, there are no ground truth labels available for recidivism of people who were never released by the judges. In this paper, we focus only on the direct influence of human feedback on data fairness i.e. the case in which humans assign labels for data.

Crowdsourcing is increasingly used to collect training data labels. Inevitably, crowdworkers have different biases, which are then reflected in the labels collected from the workers. A very recent study [8] conducted on Amazon Mechanical Turk showed that the crowdworkers were equally racially biased as COMPAS in predicting recidivism. The difference in false positive rates of crowd predictions for white and black defendants was significant and nearly equal to that of the predictions made by COMPAS. The same was true for false negative rates also. The bias didn't change much even when the crowdworkers were not explicitly displayed the race of the defendants.

We consider settings similar to [8]. Workers are asked to provide their answers (or labels) about some tasks with unknown ground truth labels. Every task has some non-sensitive details that are shown to the workers and a sensitive attribute (for example, race) that is not explicitly shown. But the sensitive attribute may potentially be correlated with the non-sensitive task details. A worker inspects the tasks assigned to her and submits labels for the tasks. Each task is assumed to have a ground truth label but the workers don't have any way of accessing the ground truth. They can only use the task details, their prior knowledge and incomplete information from other sources to make an "educated guess" about the ground truth. The examples of such tasks are "Will a defendant with given personal history recidivate within the next two years or not?" or "Will a candidate with given CV be successful in the job applied for?" or "Is given political news item fake?". The sensitive attributes in these example tasks are race, gender and political group respectively. Every worker charges a fee for answering the assigned tasks. The requester has a budget constraint on the fees that she can pay to the workers. In this paper, we make the following contributions:

(1) We propose a novel algorithm for assigning tasks to workers, which optimizes the expected accuracy of labels obtained from crowd while ensuring that the collected labels satisfy desired notions of *error fairness*. The algorithm also ensures diversity of responses by limiting the probability of assigning

AIES '19, January 27–28, 2019, Honolulu, HI, USA

ACM ISBN 978-1-4503-6324-2/19/01...$15.00
https://doi.org/10.1145/3306618.3314282

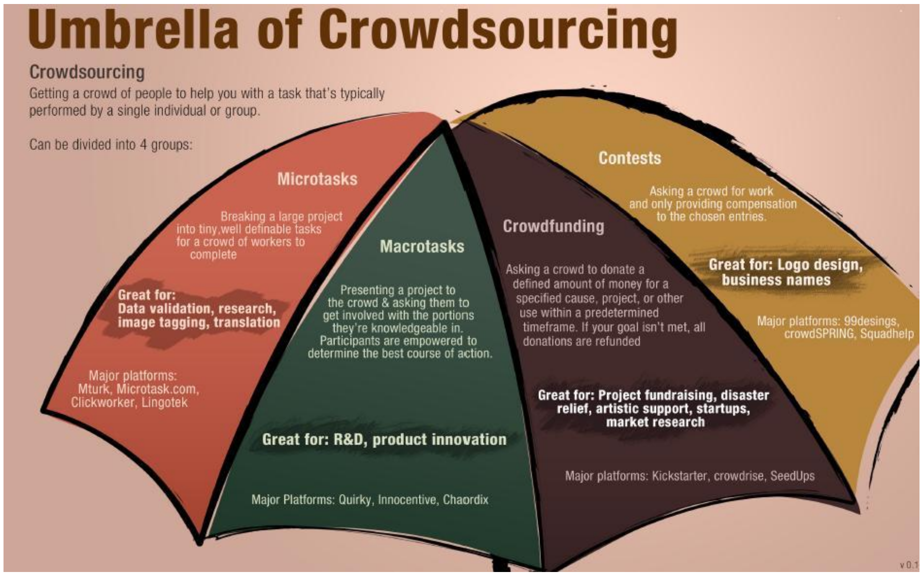
PROBLEM STATEMENT

Generally, when a task is assigned to a crowd workers. They can be discriminatory with respect to sensitive attributes such as Gender,Race,Caste, Religion etc. This raises question on the fairness of crowdsourced data for further use, such as traning a machine learning algorithm and reduce its accuracy. Moreover, This paper address the problem of unfair dataset based on linear programming and probability distribution.

Introduction of Crowdsourcing and Fairness

What is Crowdsourcing?

Crowdsourcing involves obtaining work, information, or opinions from a large group of people who submit their data via the Internet, social media, and smartphone apps. People involved in crowdsourcing sometimes work as paid freelancers, while others perform small tasks voluntarily. For example, traffic apps like Waze encourage drivers to report accidents and other roadway incidents to provide real-time, updated information to app users.



Crowdsourcing is the collection of information, opinions, or work from a group of people, usually sourced via the Internet.

- Crowdsourcing work allows companies to save time and money while tapping into people with different skills or thoughts from all over the world.
- While crowdsourcing seeks information or work product, crowdfunding seeks money to support individuals, charities, or startup companies.

- The advantages of crowdsourcing include cost savings, speed, and the ability to work with people who have skills that an in-house team may not have.

Advantages and Disadvantages of Crowdsourcing

The advantages of crowdsourcing include cost savings, speed, and the ability to work with people who have skills that an in-house team may not have. If a task typically takes one employee a week to perform, a business can cut the turnaround time to a matter of hours by breaking the job up into many smaller parts and giving those segments to a crowd of workers.

Many types of jobs can be crowdsourced, including website creation and transcription. Companies that want to design new products often turn to the crowd for opinions. Rather than rely on small focus groups, companies can reach millions of consumers through social media, ensuring that the business obtains opinions from a variety of cultural and socioeconomic backgrounds. Oftentimes, consumer-oriented companies also benefit from getting a better gauge of their audience and creating more engagement or loyalty.

But that being said, crowdsourcing isn't a magic bullet for companies that hope to lighten their workload while pursuing the next shining star of an idea. Many times, someone will have to sift through all the ideas being pitched, fundraising goals can fall short in all-or-nothing type funding platforms, and the right crowd can be difficult to find or engage.

Examples of Crowdsourcing

Companies that need some jobs done only on occasions, such as coding or graphic design, can crowdsource those tasks and avoid the expense of a full-time in-house employee.

While crowdsourcing often involves breaking up a big job, businesses sometimes use crowdsourcing to assess how multiple people perform at the same job. For instance, if a company wants a new logo, it can have dozens of graphic designers assemble samples for a small fee. The company can then pick a favorite and pay for a more complete logo package.

Fairness in Machine Learning

Machine learning based systems are reaching society at large and in many aspects of everyday life. This phenomenon has been accompanied by concerns about the ethical issues that may arise from the adoption of these technologies. ML fairness is a recently established area of machine learning that studies how to ensure that biases in the data and model inaccuracies do not lead to models that treat individuals unfavorably on the basis of characteristics such as e.g. race, gender, disabilities, and sexual or political orientation. In this manuscript, we discuss some of the limitations present in the current reasoning about fairness and in methods that deal with it, and describe some work done by the authors to address them. More specifically, we show how causal Bayesian networks can play an important role to reason about and deal with fairness, especially in complex unfairness scenarios. We describe how optimal transport theory can be used to develop methods that impose constraints on the full shapes of distributions corresponding to different sensitive attributes, overcoming the limitation of most approaches that approximate fairness desiderata by imposing constraints on the lower order moments or other functions of those distributions. We present a unified framework that encompasses methods that can deal with different settings and fairness criteria, and that enjoys strong theoretical guarantees. We introduce an approach to learn fair representations that can generalize to unseen tasks. Finally, we describe a technique that accounts for legal restrictions about the use of sensitive attributes.

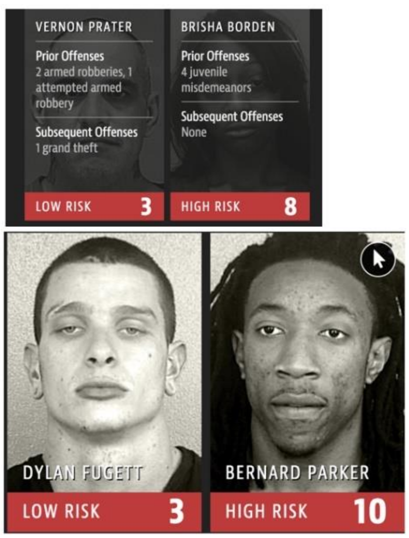
Introduction of Crowdsourcing and Datafairness

Algorithmic decision-making is gaining popularity in many diverse application areas of social importance. Examples include criminal recidivism prediction, stop-and-frisk programs, university admissions, bank loan decisions, screening job candidates, fake news control, information filtering(personalization) and search engine rankings etc. Recently, questions were raised about the fairness of these algorithms. An investigation [26] found COMPAS (a popular software used by courts to predict criminal recidivism risk) racially discriminatory. Other software systems have also been found to be biased against people of different races, genders and political views. This has led to a widespread and legitimate concern about the potential negative influence of such systems on the society. One of the main reasons of algorithmic bias is the bias in the training datasets. In order to achieve algorithmic fairness, the issue of data fairness needs to be addressed first. In many interesting cases, data is directly or indirectly influenced by some kind of human feedback. The influence is obvious and direct if human assigned labels are used as a proxy for ground truth labels. However, human feedback can also indirectly influence the so-called “ground-truth" datasets (when the labels are not human assigned but observed in reality). This is because the ground truth labels can only be collected for a finite number of data points and the selection of data points is often influenced by humans. For example, there are no ground truth labels available for recidivism of people who were never released by the judges. In this paper, we focus only on the direct influence of human feedback on data fairness i.e. the case in which humans assign labels for data. Crowdsourcing is increasingly used to collect training data labels. Inevitably, crowdworkers have different biases, which are then reflected in the labels collected from the workers. A very recent study conducted on Amazon Mechanical Turk showed that the crowdworkers were equally racially biased as COMPAS in predicting recidivism. The difference in false positive rates of crowd predictions for white and black defendants was significant and nearly equal to that of the predictions made by COMPAS. The same was true for false negative rates also. The bias didn’t change much even when the crowdworkers were not explicitly displayed the race of the defendants. We consider settings similar to [8]. Workers are asked to provide their answers (or labels) about some tasks with unknown ground truth labels. Every task has some non-sensitive details that are shown to the workers and a sensitive attribute (for example, race) that is not explicitly shown. But the sensitive attribute may potentially be correlated with the non-sensitive task details. A worker inspects the tasks assigned to her and submits labels for the tasks. Each task is assumed to have a ground truth label but the workers don't have any way of accessing the ground truth. They can only use the task details, their prior knowledge and incomplete information from other sources to make an "educated guess" about the ground truth. The examples of such tasks are "Will a defendant with given personal history recidivate within the next two years or not?" or "Will a candidate with given CV be successful in the job applied for?" or "Is given political news item fake?". The sensitive attributes in these example tasks are race, gender and political group respectively. Every worker charges a fee for answering the assigned tasks.

Compass Software

Correctional Offender Management Profiling for Alternative Sanctions (COMPAS) is a case management and decision support tool developed and owned by Northpointe (now Equivant) used by U.S. courts to assess the likelihood of a defendant becoming a recidivist.

COMPAS has been used by the U.S. states of New York, Wisconsin, California, Florida's Broward County, and other jurisdictions.



Risk Assessment

The COMPAS software uses an algorithm to assess potential recidivism risk. Northpointe created risk scales for general and violent recidivism, and for pretrial misconduct. According to the COMPAS Practitioner's Guide, the scales were designed using behavioral and psychological constructs "of very high relevance to recidivism and criminal careers.

- **Pretrial Release Risk scale**: Pretrial risk is a measure of the potential for an individual to fail to appear and/or to commit new felonies while on release. According to the research that informed the creation of the scale, "current charges, pending charges, prior arrest history, previous pretrial failure, residential stability, employment status, community ties, and substance abuse" are the most significant indicators affecting pretrial risk scores.
- **General Recidivism scale**: The General Recidivism scale is designed to predict new offenses upon release, and after the COMPAS assessment is given. The scale uses an individual's criminal history and associates, drug involvement, and indications of juvenile delinquency.
- **Violent Recidivism scale**: The Violent Recidivism score is meant to predict violent offenses following release. The scale uses data or indicators that include a person's "history of violence, history of non-compliance, vocational/educational problems, the person's age-at-intake and the person's age-at-first-arrest."

Critiques and legal rulings

In July 2016, the Wisconsin Supreme Court ruled that COMPAS risk scores can be considered by judges during sentencing, but there must be warnings given to the scores to represent the tool's "limitations and cautions."

A general critique of the use of proprietary software such COMPAS is that since the algorithms it uses are trade secrets, they cannot be examined by the public and affected parties which may be a violation of due process. Additionally, simple, transparent and more interpretable algorithms (such as linear regression) have been shown to perform predictions approximately as well as the COMPAS algorithm.

Another general criticism of machine-learning based algorithms is since they are data-dependent if the data are biased, the software will likely yield biased results.

Accuracy

In 2016, Julia Angwin was co-author of a ProPublica investigation of the algorithm. The team found that “blacks are almost twice as likely as whites to be labeled a higher risk but not actually re-offend,” whereas COMPAS “makes the opposite mistake among whites: They are much more likely than blacks to be labeled lower-risk but go on to commit other crimes.” They also found that only 20 percent of people predicted to commit violent crimes actually went on to do so.

In a letter, Northpointe criticized ProPublica’s methodology and stated that: “[The company] does not agree that the results of your analysis, or the claims being made based upon that analysis, are correct or that they accurately reflect the outcomes from the application of the model.”

Another team at the Community Resources for Justice, a criminal justice think tank, published a rebuttal of the investigation's findings. Among several objections, the CRJ rebuttal concluded that the Propublica's results: "contradict several comprehensive existing studies concluding that actuarial risk can be predicted free of racial and/or gender bias."

A subsequent study has shown that COMPAS software is more accurate than individuals with little or no criminal justice expertise and less accurate than groups of individuals. They found that: "On average, they got the right answer 63 percent of their time, and the group’s accuracy rose to 67 percent if their answers were pooled. COMPAS, by contrast, has an accuracy of 65 percent."

Introduction to Machine Learning

Machine learning is a branch of artificial intelligence (AI) and computer science which focuses on the use of data and algorithms to imitate the way that humans learn, gradually improving its accuracy.

Machine learning is an important component of the growing field of data science. Through the use of statistical methods, algorithms are trained to make classifications or predictions, uncovering key insights within data mining projects. These insights subsequently drive decision making within applications and businesses, ideally impacting key growth metrics. As big data continues to expand and grow, the market demand for data scientists will increase, requiring them to assist in the identification of the most relevant business questions and subsequently the data to answer them.

Working of Machine Learning

1. A Decision Process: In general, machine learning algorithms are used to make a prediction or classification. Based on some input data, which can be labelled or unlabeled, your algorithm will produce an estimate about a pattern in the data.
2. An Error Function: An error function serves to evaluate the prediction of the model. If there are known examples, an error function can make a comparison to assess the accuracy of the model.
3. An Model Optimization Process: If the model can fit better to the data points in the training set, then weights are adjusted to reduce the discrepancy between the known example and the model estimate. The algorithm will repeat this evaluate and optimize process, updating weights autonomously until a threshold of accuracy has been met.

Machine learning methods

Machine learning classifiers fall into three primary categories.

Supervised machine learning

Supervised learning, also known as supervised machine learning, is defined by its use of labeled datasets to train algorithms that to classify data or predict outcomes accurately. As input data is fed into the model, it adjusts its weights until the model has been fitted appropriately. This occurs as part of the cross validation process to ensure that the model avoids overfitting or underfitting. Supervised learning helps organizations solve for a variety of real-world problems at scale, such as classifying spam in a separate folder from your inbox. Some methods used in supervised learning include neural networks, naïve bayes, linear regression, logistic regression, random forest, support vector machine (SVM), and more.

Unsupervised machine learning

Unsupervised learning, also known as unsupervised machine learning, uses machine learning algorithms to analyze and cluster unlabeled datasets. These algorithms discover hidden patterns or data groupings without the need for human intervention. Its ability to discover similarities and differences in information make it the ideal solution for exploratory data analysis, cross-selling strategies, customer segmentation, image and pattern recognition. It’s also used to reduce the number of features in a model through the process of dimensionality reduction; principal component analysis (PCA) and singular value decomposition (SVD) are two common approaches for this. Other algorithms used in unsupervised learning include neural networks, k-means clustering, probabilistic clustering methods, and more.

Semi-supervised learning

Semi-supervised learning offers a happy medium between supervised and unsupervised learning. During training, it uses a smaller labeled data set to guide classification and feature extraction from a larger, unlabeled data set. Semi-supervised learning can solve the problem of having not enough labeled data (or not being able to afford to label enough data) to train a supervised learning algorithm.

Reinforcement machine learning

Reinforcement machine learning is a behavioral machine learning model that is similar to supervised learning, but the algorithm isn't trained using sample data. This model learns as it goes by using trial and error. A sequence of successful outcomes will be reinforced to develop the best recommendation or policy for a given problem.

The IBM Watson® system that won the *Jeopardy!* challenge in 2011 makes a good example. The system used reinforcement learning to decide whether to attempt an answer (or question, as it were), which square to select on the board, and how much to wager—especially on daily doubles.

Real-world machine learning use cases

Here are just a few examples of machine learning you might encounter every day:

Speech recognition: It is also known as automatic speech recognition (ASR), computer speech recognition, or speech-to-text, and it is a capability which uses natural language processing (NLP) to process human speech into a written format. Many mobile devices incorporate speech recognition into their systems to conduct voice search—e.g. Siri—or provide more accessibility around texting.

Customer service: Online chatbots are replacing human agents along the customer journey. They answer frequently asked questions (FAQs) around topics, like shipping, or provide personalized advice, cross-selling products or suggesting sizes for users, changing the way we think about customer engagement across websites and social media platforms. Examples include messaging bots on e-commerce sites with virtual agents, messaging apps, such as Slack and Facebook Messenger, and tasks usually done by virtual assistants and voice assistants.

Computer vision: This AI technology enables computers and systems to derive meaningful information from digital images, videos and other visual inputs, and based on those inputs, it can take action. This ability to provide recommendations distinguishes it from image recognition tasks. Powered by convolutional neural networks, computer vision has applications within photo tagging in social media, radiology imaging in healthcare, and self-driving cars within the automotive industry.

Recommendation engines: Using past consumption behavior data, AI algorithms can help to discover data trends that can be used to develop more effective cross-selling strategies. This is used to make relevant add-on recommendations to customers during the checkout process for online retailers.

Automated stock trading: Designed to optimize stock portfolios, AI-driven high-frequency trading platforms make thousands or even millions of trades per day without human intervention.

Implementation of Adults Dataset

Adults Dataset

An individual’s annual income results from various factors. Intuitively, it is influenced by the individual’s education level, age, gender, occupation, and etc.

This is a widely cited KNN dataset. I encountered it during my course, and I wish to share it here because it is a good starter example for data pre-processing and machine learning practices.

Fields
The dataset contains 16 columns
Target filed: Income
-- The income is divide into two classes: <=50K and >50K
Number of attributes: 14
-- These are the demographics and other features to describe a person

We can explore the possibility in predicting income level based on the individual’s personal information.

	age	workclass	fnlwgt	education	educational-num	marital-status	occupation	relationship	race	gender	capital-gain	capital-loss	hours-per-week	native-country	income
0	25	Private	226802	11th	7	Never-married	Machine-op-inspct	Own-child	Black	Male	0	0	40	United-States	<=50K
1	38	Private	89814	HS-grad	9	Married-civ-spouse	Farming-fishing	Husband	White	Male	0	0	50	United-States	<=50K
2	28	Local-gov	336951	Assoc-acdm	12	Married-civ-spouse	Protective-serv	Husband	White	Male	0	0	40	United-States	>50K
3	44	Private	160323	Some-college	10	Married-civ-spouse	Machine-op-inspct	Husband	Black	Male	7688	0	40	United-States	>50K
4	18	Private	103497	Some-college	10	Never-married	Prof-specialty	Own-child	White	Female	0	0	30	United-States	<=50K
5	34	Private	198693	10th	6	Never-married	Other-service	Not-in-family	White	Male	0	0	30	United-States	<=50K
6	29	Private	227026	HS-grad	9	Never-married	Prof-specialty	Unmarried	Black	Male	0	0	40	United-States	<=50K

Data Preprocessing

Data preprocessing is a process of preparing the raw data and making it suitable for a machine learning model. It is the first and crucial step while creating a machine learning model.

When creating a machine learning project, it is not always a case that we come across the clean and formatted data. And while doing any operation with data, it is mandatory to clean it and put in a formatted way. So for this, we use data preprocessing task.

A real-world data generally contains noises, missing values, and maybe in an unusable format which cannot be directly used for machine learning models. Data preprocessing is required tasks for cleaning the data and making it suitable for a machine learning model which also increases the accuracy and efficiency of a machine learning model.

It involves below steps:

- Getting the dataset
- Importing libraries
- Importing datasets
- Finding Missing Data
- Encoding Categorical Data
- Splitting dataset into training and test set
- Feature scaling

	age	workclass	fnlwgt	education	educational-num	marital-status	occupation	relationship	race	gender	capital-gain	capital-loss	hours-per-week	native-country	income
0	8	3	19329	4	6	1	6	3	2	1	0	0	39	38	0
1	21	3	4212	2	8	0	4	0	4	1	0	0	49	38	0
2	11	1	25340	3	11	0	10	0	4	1	0	0	39	38	1
3	27	3	11201	3	9	0	6	0	2	1	98	0	39	38	1
4	1	3	5411	3	9	1	9	3	4	0	0	0	29	38	0

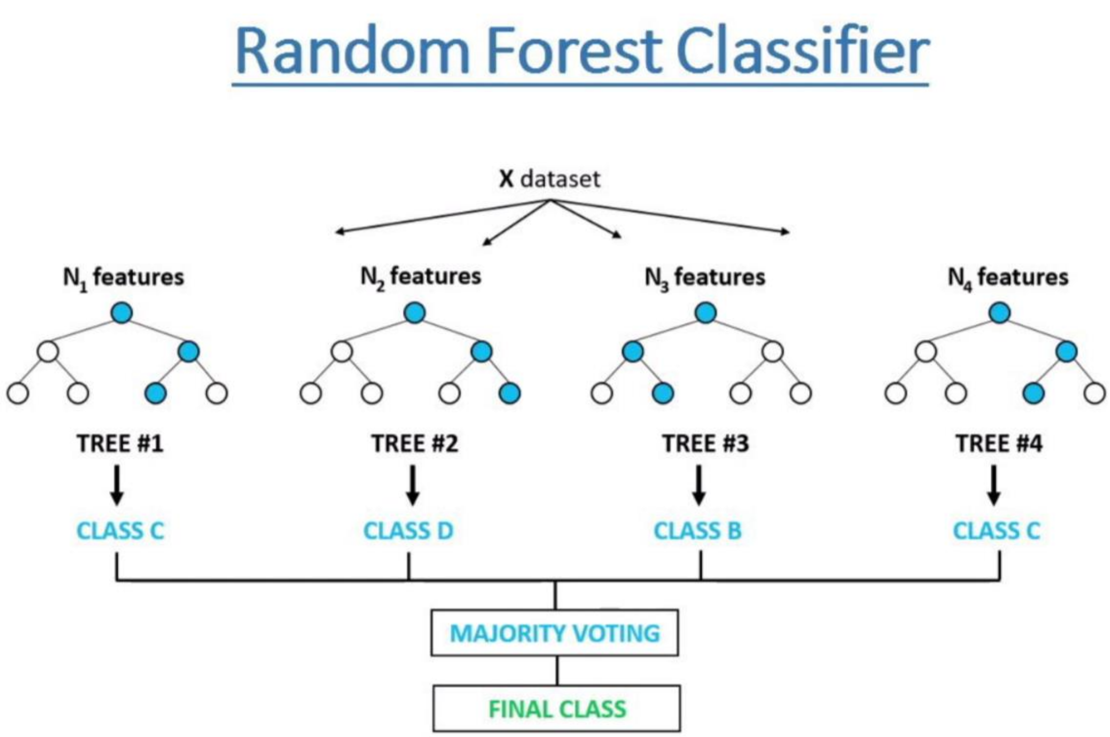
Random Forest Classifier

Random Forest is a popular machine learning algorithm that belongs to the supervised learning technique. It can be used for both Classification and Regression problems in ML. It is based on the concept of ensemble learning, which is a process of *combining multiple classifiers to solve a complex problem and to improve the performance of the model.*

As the name suggests, *"Random Forest is a classifier that contains a number of decision trees on various subsets of the given dataset and takes the average to improve the predictive accuracy of that dataset."* Instead of relying on one decision tree, the random forest takes the prediction from each tree and based on the majority votes of predictions, and it predicts the final output.

The greater number of trees in the forest leads to higher accuracy and prevents the problem of overfitting.

The below diagram explains the working of the Random Forest algorithm:



Below are some points that explain why we should use the Random Forest algorithm:

- It takes less training time as compared to other algorithms.
- It predicts output with high accuracy, even for the large dataset it runs efficiently.
- It can also maintain accuracy when a large proportion of data is missing.

Applications of Random Forest

There are mainly four sectors where Random forest mostly used:

1. Banking: Banking sector mostly uses this algorithm for the identification of loan risk.
2. Medicine: With the help of this algorithm, disease trends and risks of the disease can be identified.
3. Land Use: We can identify the areas of similar land use by this algorithm.
4. Marketing: Marketing trends can be identified using this algorithm.

Advantages of Random Forest

- Random Forest is capable of performing both Classification and Regression tasks.
- It is capable of handling large datasets with high dimensionality.
- It enhances the accuracy of the model and prevents the overfitting issue.

Model Fitting and Accuracy result

Now we will fit the Random forest algorithm to the training set. To fit it, we will import the Random Forest Classifier class from the sklearn.ensemble library. The code is given below:

```
from sklearn.ensemble import RandomForestClassifier

rfc = RandomForestClassifier()

model1 = rfc.fit(X_train, y_train)
prediction1 = model1.predict(X_test)

print("Acc on training data: {:,.3f}".format(rfc.score(X_train, y_train)))
print("Acc on test data: {:,.3f}".format(rfc.score(X_test, y_test)))
```

```
Acc on training data: 1.000
Acc on test data: 0.857
```

```
from sklearn.metrics import confusion_matrix
from sklearn.metrics import classification_report
```

```
print(confusion_matrix(y_test, prediction1))
```

```
[[10366   772]
 [ 1317  2198]]
```

```
print(classification_report(y_test, prediction1))
```

```
              precision    recall  f1-score   support

           0       0.89      0.93      0.91     11138
           1       0.74      0.63      0.68      3515

    accuracy                           0.86     14653
   macro avg       0.81      0.78      0.79     14653
weighted avg       0.85      0.86      0.85     14653
```

```
from sklearn.metrics import accuracy_score
accuracy_score(y_test, prediction1)
```

```
0.8574353374735549
```

Fair Classification and Final Output

GENDER(sensitive attribute) with its actual and predicted values

	Gender	Y_test	Prediction
38113	1	0	0
39214	1	1	0
44248	1	1	0
10283	1	0	0
26724	0	0	0
...	...	...	...
48826	0	1	0
44230	0	0	0
27824	1	0	0
13582	1	0	0
14557	1	0	0

Count of GENDER(sensitive attributes) with its accuracy of prediction

```
6780
For Gender = 1 is Correct Accuracy is 0.6963845521774856
```

```
2956
For Gender = 1 is InCorrect Accuracy is 0.3036154478225144
```

```
4358
For Gender = 0 is Correct Accuracy is 0.44761709120788823
```

```
559
For Gender = 0 is InCorrect Accuracy is 0.057415776499589156
```

Conclusion

This research paper address the problem of data fairness in crowdsourcing with the help of novel crowdsourcing algorithm that learns an optimal selection on probability distribution over the available set of workers to maximize the expected accuracy of collected data, while ensuring that the errors in the data are not unfairly and discriminatory towards any particular social group.

Its only drawback is it do not define data fairness for the case of subjective task which have no ground truth label.

References

- https://www.iitrpr.ac.in/
- https://archive.ics.uci.edu/ml/datasets/adult
- https://en.wikipedia.org/wiki/COMPAS_(software)
- https://www.ibm.com/cloud/learn/machine-learning
- https://www.javatpoint.com/machine-learning-random-forest-algorithm