

Emergence of Technologies – CAO Allocation

Dissatisfaction with the CAO system has led a new Minister for Education to propose that the allocation of places in University Programmes should be determined by an AI system based on a Deep Neural Network. The proposed allocation system would be based on a machine learning algorithm trained on the input data of previous students (the admissions information available in your MyTCD profile) along with the ultimate performance of these previous students. The system will be trained to take account of DARE and HEAR students. The system will also take input scores from the HPAT and the assessment of portfolios where appropriate.

The Minister has heralded the proposed system as the start of a “Brave New World” where places are allocated to the students who are most likely to benefit from being offered a place. The new system will have the flexibility to recognise that a high grade at leaving cert in French is of more significance than a high grade in Irish, for entry onto a degree in French, but because it will be trained on a large dataset of previous students it will identify the optimal weighting. Above all the Minister has promoted the new system on the basis that it will be fair.

This proposal has given rise to concerns among many stakeholder groups, including: candidates, universities, schools, parents organisations, student unions, bodies promoting equal access to third level education and employers. To reassure these groups, the Minister has promised an independent review of the proposed system and your company has been commissioned to undertake this review.

This is the first meeting of the team charged with assessing the proposed system

(1) Carry out an initial “Brainstorming” session to identify the potential risks and benefits of the new system. Your session should capture the risks and benefits of all interested parties.

The new AI-based course allocation system has the potential to assign candidates in a way which much better reflects their abilities. This has numerous benefits. Firstly, the standard of candidates will be generally higher, as the AI will prevent students who lack ability in their subject from participating in that course. This means that dropout rates will be lower, and that ultimately, people will specialise in fields they are actually competent in, with the potential to raise productivity and grow the economy in the long term. This could also be good for career satisfaction. Almost universally, when people are good at something, they enjoy it.

However, in practice, many of these goals may not come to fruition and there are likely to be serious flaws in the system if there are inadequate safeguards and insufficient human oversight of the system.

No matter what course allocation system is put in place, students and parents will attempt to exploit it. For example, they will take courses to train for aptitude tests such as the SAT or HPAT, even though these exams are not intended to be studied for. Although the aim of the system is to assign candidates to the course they are best at, they are likely to try their hardest to get into the course they choose anyway. This could be particularly problematic in the case of an AI

system, as currently, there is no way to know exactly how neural networks make their decisions - which data they take into account and how they are weighted.

Applicants will be left guessing at what actions they should take to maximise their chances of getting into highly sought after courses. Artificial intelligence as it exists today is prone to exploitation. This can be seen on social media, where people will use certain keywords and hashtags (whether or not they are relevant to the content of the post) with the aim of being amplified by the algorithms these sites use to determine which content to promote. This means that these algorithms have to constantly adapt to avoid this exploitation. A similar phenomenon could take hold under this new system. For example, students might fare better if they hyper focused on a handful of subjects, and neglected their general education.

Under the current system, the best way to get into any course is to achieve high grades in the leaving certificate, by studying hard. While there are problems with this system, it is at least transparent. If the new system is too complex, it has the potential to cause students unnecessary stress.

Furthermore, this AI selects for specialised talent over general ability and hard work. Although one could argue that there are advantages to this, we should be careful not to overly specialise, as interdisciplinary knowledge and a wide range of perspectives are valuable in any industry.

Finally, because of the inherently opaque nature of this system, it is likely to generate outrage and lack of trust from the public. If a student fails to get into a course via the CAO, they know exactly why. This is an advantage of Ireland's system compared to those of the UK or USA. This new system lacks accountability and transparency. To ameliorate these issues, the government should encourage and take advantage of efforts to develop transparent neural networks which can explain how and why they made the decisions they did.

(2) Your group is aware that previous decision-making systems have exhibited unintended bias that came from the training data. This is a particular concern for many of the stakeholders. Please indicate how the data should be preprocessed to minimise bias in explaining your choices. If you identify any potential sources of bias that cannot be removed, then please comment on them.

The data sets need to be trained to eliminate as many biases as possible. Major outliers should be removed from data inputs. Continuous checks would be necessary to ensure data is accurate.

To ensure fairness, HEAR and DARE are taken into account in the AI's system. However some students with disadvantaged backgrounds or disabilities do not qualify for HEAR and DARE. There is a risk of biased data for students with late diagnosed disability who do not qualify for DARE. Students with disabilities can still be entitled to certain school aids (such as reasonable accommodations) even if they don't meet DARE criteria. It should be possible for students with late-diagnosed disabilities to apply to have their past grades adjusted as though they had

Group 5 Part 1

reasonable accommodations during that period. This is to ensure fairness amongst those entitled to reasonable accommodations but who may not have had them reflected in past grades.

Private information should not skew the data. Students should be kept as anonymous as possible. Numbers should continue to be used as student identifiers as opposed to names. Even when AI is not given direct access to demographic data it can sometimes infer these characteristics through other data. For example, the AI may be able to predict the gender or nationality of a candidate from their subject choice. It is vital that personal demographics are not used as indicators.

New college courses would have a lack of data. College lecturers and course coordinators running new courses are needed to help guide AI with allocation into these new courses. Indicators of good student performance would be determined by these teachers and administrators. Similar courses and their corresponding entry requirements should also be used as data indicators.

Unfortunately there are some sources of bias that cannot be removed or accounted for. Shareholders must be made aware of this. There are intrinsic faults that lie within the current education system and these would be carried over, if not amplified, by the new AI system.

Affluence affects education. The family wealth you are born into has a profound effect on your education. Grades differ between DEIS, public and private schools, not because of individual student intelligence, but because of unequal access to resources. Those who grow up in socio-economic areas of disadvantage or in low income households are far less likely to have access to private education or extra tuition. DEIS, public and private schools differ in the learning resources and facilities they can provide for their students. More affluent students receive more educational support and are more likely to achieve higher grades. The AI is basing college allocations on grades alone, similar to the current Leaving Certificate. Grades are not necessarily a measure of intelligence and are biased towards students who receive more education support. This bias cannot be removed or accounted for without complete reform.

Gender biases are also deeply rooted in our education system. Historically Ireland has seen the gendered division of education with all-girls and all-boys schools. This gender division continues today and affects the subjects available to students. All girls schools are more likely to offer Home Economics and Art to students, while all-boys schools are more likely to offer Woodwork and Technical Graphics. This immediately puts students at a gendered disadvantage in the AI system if they want to study a subject that was not offered to them in school. This issue also expands beyond gender, and affects those in smaller schools where not as many subjects may be offered. If the AI uses optimal weighting of subject grades to allocate college places, it is will contain gender biases and discriminate against those who are unable to sit the optimally weighted subjects that align with the course they want.

(3) Design a series of tests that could be carried out on the new system, post-training, to check that it is functioning properly and does not display bias.

Any biases in the training data will be reflected in the output. For example, if you were trying to predict who could perform well in a certain course, Group A might have historically been performing better than group B. The neural network might just assume that any student from group A might just perform better than group B. Which wouldn't be unfounded but still would be quite unfair. To prevent this we would need to look at all the data provided and judge if there are systemic biases in the data, figure out why they exist and test if they still exist.

One method that one could use is to make 10000 applications with one specific group that historically did better. Isolate all other factors and see if the bias persists in other words, give the neural network 5000 applications in which the only difference is for example race/ethnicity keeping all other variables the same. Will the neural network unfairly discriminate against the students based on their ethnicity? If the answer is yes then we tweak the system and assess the training data and if necessary completely retrain the neural network.

One important thing to mention is that even if you try not to let it have access to some information it might still be able to work out the biases. An example of this would be a bank making an AI predict whether someone would default on a loan but didn't give the AI the race (this is often a protected status) but it still learned to guess the race based on where they lived.

Race is only one of the domains at which comparisons could be made, it could be gender, sexual orientation, socio-economic background or any other group that could infer biases.

As a result, the AI needs to fall under the scrutiny of multiple tests with different scenarios and humans need to decide if the outcomes are fair by isolating certain behaviours. This needs to be a transparent process so that all parties can see and give feedback.

As a result, the system needs to fall under scrutiny multiple times and then after many cycles and iterations of public scrutiny and rigorous "fairness testing". Only then can we have a deep neural network make course allocation decisions.

Group 5 Part 1

Group 5 Member (Names and Student Numbers)

Name	Student Number
Anna Coghlan	21365544
Emma Stewart	21366848
Seamus O'Sullivan	20333628
Arshad Reham Mohammed	20334119
Kenneth Johnson	18310776
John Ryan	21363606
Dishant Tekwani	20309212
Malachy Solan	21363589