The Journal 2022

Editorial

American Theologian, Joseph Fletcher, understood the word 'conscience' not as a noun: not as something we have, or use, but rather something we do. Deciding what is right or wrong is not something that is predetermined for us, but is rather a process of discovery based on what is in front of us. I see 'scholarship' in a similar light. Whilst the giving of scholarships is something that we do at the RGS, our school value of Scholarship is also the process of discovery: nurturing a love of learning, cultivating good learning habits, and fostering a scholarly mindset. In other words, *scholarship for all* is genuinely something that is cultivated within the classrooms at the RGS, and this volume of The Journal is a celebration of the quirky, opinionated, passionate, and remarkable students we are so lucky to call members of our community.

In a year of recovery and healing, the members of the Upper Sixth were not tasked with completing an *Independent Learning Assignment* (as is the norm), but were instead given the opportunity to flex their academic muscles in other ways, through external essay competitions and internal scholarship opportunities, ORIS projects and fieldwork. The writing within this compendium was completed as a result of something that cannot be examined nor quantified, but rather comes from a pure love of learning and discovery. That is what makes the RGS such a special place: students of all disciplines, all ages, show a real yearning for knowledge for its own sake, and this edition of *The Journal* is their academic fix.

I hope you enjoy reading the varied and imaginative works contained in these pages. They include a frank assessment of the use of electric vehicles, a discussion on conflict in developing countries, an analysis of machine learning for sparse data, the invention of a brand new language, and much more. Perhaps you will be inspired by our students' work – they certainly inspire me on a daily basis.

Continuing the remarkable work of my predecessor, Mr Bradford, has not been a simple task. The legacy of *Scholarship for All* he has left behind is, quite honestly, something that has fundamentally improved the ethos of the school, and is part of our everyday lives as members of the RGS community. His aim was to stoke the fire of scholarship in the school, and he has left a blazing fire of academic enrichment. I wish him all the very best in his promising career.

Mr S Herman-Wilson
Head of Scholarship

Contents

All references to appendices have been redacted from this publication but are available on request.

Why do developing countries suffer more frequently from conflict than the developed countries?

Entry to the R.A. Butler Politics Essay Prize, 2021

Alex Mitchell, Upper Sixth

Conflict – the action of opposing groups competing with one another by means of violence in pursuit of mutually incompatible goals – has throughout human history been a tragically needless yet consistently pervasive threat to life and wellbeing. In modern times, while international players and factors continue to play an undeniable role, the vast majority of conflict is intra-state – it takes place between actors within a single state. Conflict is not spread evenly across the world; it is, overwhelmingly, centred in developing countries. While the consistency with which this is observed means the connection is for many almost an unconscious assumption, examining why this connection exists is crucial. If we can determine what exactly makes developing countries more likely to encounter conflict, it brings us closer to identifying the responses – whether from local governments, developed countries or international organisations – necessary to sustain a higher degree of peace. This essay will evaluate the different ways in which underdevelopment can be linked to the prevalence of conflict, and attempt to reach a general explanation as to why this connection is so consistently observed.

The most simple and immediate explanation for the connection is the incentive of material gain – there is more conflict in developing countries because people are generally much poorer, and hence more inclined to turn to violent conflict as a means to obtain wealth or resources. It is certainly true that there is a correlation between lower income and a higher prevalence of conflict – a 2003 study found that a $1000 drop in average income resulted in a 41% increase in the chance of civil conflict (Fearon & Laitin, 2003). In an area with a weak economy, and a scarcity of alternative opportunities, it makes sense that many would turn to war as a means to increase their personal wealth. This wealth can be obtained through looting, as well as the illicit trading of arms and natural resources. Natural resources, in particular, have often been observed as a motivator for conflict in developing countries – which often, paradoxically, are rich in them. For example, studies have found that the discovery of offshore oil – where it is impossible to access without existing infrastructure – has little impact on conflict, while the discovery of onshore oil, especially in regions with incomes lower than the national average, significantly increases conflict risk (Ross, 2015).

It is also certainly true that the connection also exists in the other direction: conflict itself harms a country's development. Experience of war has been found to have significant negative effects in areas such as education, GDP per capita and undernourishment (Gates et al, 2012). A 2016 study which examined conflict in Africa by dividing it into a grid of small cells found that growth is reduced by 4.4% for each year in which a cell experiences over 50 fatalities (Ray & Esteban, 2017 cited Mueller, 2016). Specific evidence of the damaging effect of conflict on development can be found in a comparison between Burkina Faso and Burundi; both were experiencing the same steady GDP growth until Burundi's civil war in the 1990s, which saw a collapse in growth from which it had not recovered even by the 2010s (Gates et al, 2015). This is unsurprising – conflict is extremely disruptive, and makes it impossible for a local economy to properly function. Additionally, it becomes easier to mobilise a group of people to fight a war when collective memories of prior conflict can be called upon (Stewart, 2002). This suggests that the connection is self-perpetuating - underdevelopment leads to conflict, which in turn further harms development.

However, this cycle comes far from providing an explanation to why developing countries are more prone to conflict. The first problem is that the simple 'material gain' hypothesis is flawed – groups lacking wealth or resources, in a country lacking wealth or resources, would have neither the means nor the incentive to initiate a war. This is where the importance of one crucial factor becomes clear: inequality. This is not inequality of the type typically seen in developed countries (large gaps between the richest and poorest individuals) – it is inequality between different groups, usually defined along ethnic or religious lines. It is these disparities which often fuel conflict – groups which are systemically disadvantaged compared to other groups within the same state have both the means and the motivation to wage a civil conflict. They may not necessarily be in abject poverty (hence the means), but the relative reward of continuing normal economic activity from a disadvantaged position, compared to seizing resources from more affluent groups, is such that there is a clear economic incentive for conflict - hence the motivation (Ray & Esteban, 2017). These economic disparities are inextricably connected with the ethnic or religious tensions between different groups. Examples of conflicts which have emerged in such a way are numerous: the Rwandan Civil War was a result of tension between ethnic Tutsis and Hutus, and the Sri Lankan Civil War was a result of systemic discrimination against the Tamil population by the Sinhalese government. The discrimination suffered in these examples went beyond economic disadvantage – the Tutsis and Tamils had suffered exile and active persecution at the hands of the respective national governments. However, there need not even be a truly significant economic difference to stoke conflict – tension between groups can be such that a mere improvement in the economic fortunes of a rival group can provoke violence. This has been observed in the city of Meerut in India, where increases in the per capita income of the Muslim minority consistently coincided with an increase in violent conflict (Ray & Esteban, 2017 cited Mitra & Ray, 2014).

The second problem with the cycle is that it does not explain why tensions are more likely to escalate into conflict in developing countries than in developed ones. Internal disquiet is not exclusive to developing countries; there are major independence movements for several Spanish regions, and there has been deep historic unease between the Catholic and Protestant populations in Northern Ireland. While these tensions have spilled over into incidents of violence, there has never been any real concern of a full-scale conflict. One crucial reason for this is a feature which developed countries have, but developing countries lack: strong institutions. In a developed country, there is usually a state which is strong while also being stable and accountable, and is hence able to manage tensions and maintain a high standard of peace in its territory. States in developing

countries often have much looser control over their territory and much less accountability – this means they are far less effective at managing conflict, and more likely to actually provoke it in the first place.

Arvind Ganesan and Alex Vines argue that the significance of 'greed' for resources as a motivator for rebel groups to initiate conflict has been exaggerated, and more importance should be placed on the role that a dysfunctional and unaccountable state plays in mismanaging these resources. The Angolan Civil War serves as a key example of how conflicts can be misinterpreted as a struggle for resources. Whilst the UNITA rebel group was earning around $300 million in diamond sales from 1999-2002, the conflict began in 1975, and it was not until 1994 – when the US and South Africa cut off financial assistance – that they began financing themselves in this way. The sale of diamonds was a means to finance the conflict, rather than an end in and of itself. This was further underlined when UNITA agreed to a ceasefire after the death of its leader Jonas Savimbi in 2002 despite still standing to make $1 million a day from diamonds (Ganesan & Vines, 2004). Instead, Ganesan and Vines propose multiple ways in which the combination of resources and dysfunctional states has consistently provoked conflict. The determination with which unaccountable governments hoard and exploit resources, including siphoning off revenue to line the pockets of wealthy elites, means the wider population see little benefit, creating a poorer and more resentful environment which is likely to foment rebellion. Furthermore, these governments often utilise their resource income to fund further conflicts; unlike in a developed country with strong democratic institutions, popular opposition to the waging of unnecessary wars is a non-factor. For example, Liberia's Charles Taylor used profits from shipping and timber to fund extensive military action, both within Liberia as well as supporting rebels in neighbouring Sierra Leona. Finally, the combination of a weak, resource-rich state and nearby unaccountable states can result in the latter intervening in and exacerbating a civil conflict in the former in order to gain resources – as was seen during the Second Congo War, during which a UN Panel condemned Rwanda, Uganda and Zimbabwe for plundering Congolese territory (United Nations, 2001). Weak institutions make the outbreak of conflict much more likely – and once conflict does occur, these institutions are likely to be weakened further or even destroyed.

With horizontal inequality and weak institutions now factored in, the prevalence of conflict has begun to make more coherent sense. In an environment where some groups are not well-off, these groups may seek to obtain resources from more advantaged groups, or from a corrupt, resource-rich state. An unaccountable state will seek to extract resources from its population or neighbouring countries, and put these to use by funding further conflict. The dysfunctionality and weakness of states in many developing countries also means they struggle to suppress rebel movements and conflict within their own borders. However, this still fails to explain why it is that developing countries tend to have both weak institutions and competing ethnic or religious groups. The answer to this lies, at least to some degree, in an experience shared by the majority of the world's developing nations – colonisation.

It is crucial in this situation to make a distinction between settler colonialism and exploitation colonialism. Settler colonialism involved a colonial power moving immigrants from the home country into colonised territories, usually resulting in the indigenous population being all but replaced. These settlers then usually sought to establish a society which bore strong resemblance from that of their home country. The modern states which can trace their origins back to this form of colonisation – such as Australia and the United States – are likely to be developed. The purpose of exploitation colonialism, on the other hand, is to exploit; extracting resources and human labour (often in the form of slavery) for the financial benefit of the home country. This entails a much more limited colonist presence. This was the predominant form of colonialism in Africa and Asia, and the countries which have emerged from these former colonies are almost exclusively developing.

It is valid to ask if, rather than conflict and underdevelopment contributing to and exacerbating one another, both are partially a result of the colonial experience. While some scholars have suggested that the reforms brought about by colonisation actually benefited many colonised territories, these claims, which have long been deployed to justify imperialism, are largely flawed. In Africa, while the Europeans did certainly introduce some new technology and trading links, the deeply extractive nature of the relationship between coloniser and colonised meant that the vast majority of normal people saw no benefit from these changes, instead experiencing a severe deterioration in their living conditions (Heldring & Robinson, 2013). These benefits could have been established by international organisations or missionary groups in a way which genuinely benefited Africans; instead, claims of 'civilisation' largely served as a guise by which European nations stunted the development of their territories for their own gain.

Simultaneously, the foreign extraction experienced by colonised territories created an environment where violence was encouraged. The rewards garnered from normal, productive economic activity would inevitably be taken away by the colonial power, leading many to turn to violent, unproductive economic activity (stealing and looting), the value from which generally could not be taxed or extracted (Nunn, 2007). In Africa, this exacerbated a process that had already begun before direct colonisation by the transatlantic slave trade. While slave trading had already existed before the arrival of the Europeans, they completely transformed its scale and nature; unprotected villages were often raided and entire wars were conducted with the aim of capturing slaves from enemies (Mintz & McNeil, 2018). The high demand for forced labour drove local powers to transition from the export of other goods to focus more exclusively on the violent business of the capture and sale of slaves – from 1806 to 1821, slave exports increased 240%, while rice and wheat production fell by 88% and 95% respectively (Nunn, 2007 cited Austen, 1987). The brutality and unfairness of colonial rule also inevitably resulted in countless independence movements, many of which sought to achieve their goals via war. All of this combined means that formerly colonised nations have a long history of conflict – and, as we have seen previously, prior instances of conflict increase the likelihood that it will reoccur. In settler colonies – now largely developed and peaceful – a large proportion of the economic value remained in the hands of the settlers rather than the home country, there was no capture of slaves, and independence usually came about in a much more peaceful fashion. The abruptness with which many territories were decolonised, such as Britain's sudden withdrawal from India in 1947, ensured that most newly emerging nations struggled to establish the strong institutions and infrastructure to even begin to tackle these problems.

As well as fomenting violence in the societies of future developing countries, colonialism shaped these countries in a manner which encouraged intrastate conflict long after the retreat of the colonial powers. Horizontal inequality, as detailed earlier, is a major driver of conflict; countless modern states in the developing world which combine multiple religious or ethnic groups, often with disparate levels of wealth and influence, owe their situation to colonialism. (Bayeh, 2015). The Rwandan Civil War, for example, broke out as a result of tensions between Tutsis and Hutus. While colonial powers, in this case Germany and Belgium, did not create these tensions, they exacerbated them by strongly favouring the Tutsis and introducing racial identity cards, entrenching the division and disparity between the two groups. These heightened tensions set the scene for the persecution and conflict which would eventually culminate in the

Rwandan genocide. In Sudan, the domination of the richer Muslim north over the poorer Christian and Animist south created grievances which fuelled years of brutal civil war, and were only eventually resolved by South Sudan gaining independence in 2011. The decision by the British to integrate their northern and southern Sudanese colonies – which had previously been deliberately administered separately – in 1946 created an inherently unstable country which was inevitably prone to conflict. Again, the full blame cannot be placed on the coloniser; the British decision to integrate the colonies was largely due to northern Sudanese pressure, and the years of conflict were partially a result of the Sudanese government's desire to retain control of lucrative oil fields on the border with the south. Nonetheless, it was Britain's position of power and acquiescence to northern demands which enabled a state such as Sudan, with two deeply separate and unequal groups, to come into being. Borders which cut across ethnic groups are also a common point of contention in interstate conflicts; the presence of a large number of Somalis in eastern parts of Kenya has led to poor relations and regular skirmishes between the two countries.

None of this is to suggest that colonialism can be solely blamed for conflict in the developing world. Divisions and tensions between different groups of people have always been an inevitable reality, and today's developing countries certainly were not idyllic bastions of peace and harmony before the arrival of European nations. There have continued to be countless conflicts which cannot trace their origins back to colonialism in any real sense; the Iran-Iraq war, for example, is generally considered to be the result of long-term Sunni-Shi'a division, the preceding Islamic revolution in Iran and Saddam Hussein's desire to expand Iraq's regional influence (Parasiliti, 2003). The impact of colonialism is felt most keenly in Africa, and less so in other parts of the developing world. However, what is undeniable is that colonialism, through continuous economic extraction and ill-advised state-building, has had a significant effect, both by stunting development and by creating conditions which encourage conflict. If today's developing countries had been allowed to develop freely during the periods of time which they spent subjugated by European powers, it seems likely that they would have been able to make far more progress towards becoming both economically developed and peaceful.

The interference of foreign powers in developing regions certainly did not conclude after the end of colonialism. The onset of the Cold War in the 1940s and 50s initiated an ongoing trend of conflicts in the developing world, often civil wars, becoming 'internationalised', with powerful foreign nations supporting opposing sides, each hoping to advance their national interests. In the context of the link between underdevelopment and conflict, this form of intervention is critically distinct from colonialism; due to the absence of directly harmful economic extraction, international interference in internal disputes is a way in which underdevelopment can result in conflict, rather than a factor that results in both. Developing countries, with weak institutions and inherently less power and influence on the international stage, are an attractive location for powerful nations to engage in indirect conflict with one another without instigating a larger-scale international war. The US and Soviet Union regularly became involved in local wars as a proxy for fighting one another, usually seeking to establish regimes which would fall under their sphere of influence; the Vietnam War, for example, can be interpreted as a war waged by the US to stop Vietnam from becoming communist, which would not have occurred on any kind of comparable scale had they left the country to decide its own fate (Appy, 2018). Foreign interference stokes pre-existing internal tensions, encouraging and accelerating the outbreak of war – and, by means of financial and military assistance, expanding the scale of the resulting violence. The Angolan Civil War, for example, became a proxy conflict - the MPLA received support from the Soviet Union and Cuba, while UNITA was backed by the US and apartheid South Africa. The additional funding, firepower and troops from these external sources heightened the scale of the conflict.

Foreign interference of this kind certainly did not stop at the end of the Cold War; in fact, internationalised intrastate conflicts have become increasingly prevalent since 2008 (Uppsala Conflict Data Program, 2020). The ongoing rivalry between Saudi Arabia and Iran, both of which seek to wield influence over the rest of the Middle East, has been a key contributor to this. Seeking to foment opposition to Shi'ite Iran and Iranian-supported regimes, Saudi Arabia has encouraged the hardening of Sunni-Shi'a divisions by means of government-sponsored anti-Shi'ite propaganda programmes, primarily targeting education and the media. Both nations have encouraged conflict in increasingly fragile states by supporting opposing sides (Fisher, 2016). In Lebanon, Saudi and Iranian backing of the Sunni government and Shi'ite Hezbollah group respectively contributed to the escalation of a 2008 political crisis into an outbreak of violence. In Syria, Saudi Arabia has provided funding and equipment to rebels, including Sunni Islamists, while Iran has provided troops to fight on the behalf of the government. In Yemen, Saudi Arabia responded to the overthrow of the pro-Saudi government by rebels associated with Iran by conducting a brutal bombing campaign which has affected millions of civilians. All of these examples demonstrate the way that, even after the end of the Cold War, foreign intervention continues to exacerbate conflict in developing countries.

To conclude, there are three ways in which the underdevelopment of nations and their increased experience of conflict can be linked: conflict as a result of underdevelopment, underdevelopment as a result of conflict, and both as a result of a third factor (colonialism). All of them certainly play a key role in explaining why developing countries suffer more frequently from conflict. Underdevelopment can result in conflict by increasing the material incentive for violent activities, as well as by making a country more susceptible to become a battleground for larger powers. Conflict can result in underdevelopment by damaging institutions and disrupting normal economic activity. Colonialism resulted in both by inflicting major economic extraction, failing to establish strong institutions, and exacerbating disparities between ethnic groups before pushing them together into new, fragile states. The combination of these factors brings us back again to a cycle of conflict; countries which are underdeveloped or conflict-prone become trapped, each problem compounding and exacerbating the other. The colonial experience is likely to have contributed significantly to this entrapment, or stunted efforts to escape it, in many countries. This does not mean the cycle is inescapable. Today's developing countries were, in the past, underdeveloped and afflicted by war. However, any efforts to relieve conflict in a country or aid its development cannot be undertaken in isolation – both issues must be faced simultaneously, or neither will be solved.

BIBLIOGRAPHY

- Appy, C. (2018). *What Was the Vietnam War About?* [online] New York Times. Available at: https://www.nytimes.com/2018/03/26/opinion/what-was-the-vietnam-war-about.html (Accessed 25/06/21)
- Bayeh, E. (2015). The Political and Economic Legacy of Colonialism in the Post-Independence African States. *International Journal in Commerce, IT and Social Sciences*, [online] Volume 2(2), p.89-93. Available at: https://www.researchgate.net/publication/273577309_THE_POLITICAL_AND_ECONOMIC_LEGACY_OF_COLONIALISM_IN_THE_POST-INDEPENDENCE_AFRICAN_STATES (Accessed 25/04/21)

- Fearon, J. and Laitin, D. (2003). Ethnicity, Insurgency, and Civil War. *The American Political Science Review*, [online] Volume 97(1), p.75-90. Available at: https://www.jstor.org/stable/3118222 (Accessed 29/06/21)
- Fisher, M. (2016). *How the Iranian-Saudi Proxy Struggle Tore Apart the Middle East.* [online] New York Times. Available at: https://www.nytimes.com/2016/11/20/world/middleeast/iran-saudi-proxy-war.html (Accessed 14/07/21)
- Ganesan, A. and Vines, A. (2004). *Engine of War: Resources, Greed, and the Predatory State.* [online] Human Rights Watch. Available at: https://www.hrw.org/legacy/wr2k4/download/14.pdf (Accessed 19/03/21)
- Gates, S., Hegre, H., Nygård, H. and Strand, H. (2012). Development Consequences of Armed Conflict. *World Development*, [online] Volume 40(9), p.1713-22. Available at: https://havardhegre.files.wordpress.com/2013/09/development-consequences-of-armed-conflict.pdf (Accessed 07/07/21)
- Gates, S., Hegre, H., Nygård, H. and Strand, H. (2015). *The consequences of internal armed conflict for development (part 1).* [online] Stockholm International Peace Research Institute. Available at: https://www.sipri.org/commentary/blog/2015/consequences-internal-armed-conflict-development-part-1 (Accessed 08/07/21)
- Heldring, L. and Robinson, J. (2013). *Colonialism and development in Africa.* [online] VoxEU. Available at: https://voxeu.org/article/colonialism-and-development-africa (Accessed 5/07/21)
- Mintz, S. and McNeil, S. (2018). *Enslavement.* [online] Digital History. Available at: https://www.digitalhistory.uh.edu/disp_textbook.cfm?smtid=2&psid=445 (Accessed 30/06/21)
- Nunn, N. (2007). Historical legacies: A model linking Africa's past to its current underdevelopment. *Journal of Development Economics*, [online] Volume 83(1), p. 157-75. Available at: https://scholar.harvard.edu/files/nunn/files/legacy_jde.pdf (Accessed 14/04/21)
- Parasiliti, A. (2003). The Causes and Timing of Iraq's Wars: A Power Cycle Assessment. *International Political Science Review / Revue Internationale De Science Politique*, [online] Volume 24(1), p.151-65. Available at: http://www.jstor.org/stable/1601336 (Accessed 02/07/21)
- Ray, D. and Esteban, J. (2017). Conflict and Development. *Annual Review of Economics*, [online], Volume 9(1), p. 263-93. Available at: https://pages.nyu.edu/debraj/Papers/EstebanRayAnnReviews.pdf (Accessed 10/03/21)
- Ross, M. (2015). What Have We Learned about the Resource Curse? *Annual Review of Political Science*, [online] Volume 18(1), p.239-59. Available at: https://escholarship.org/uc/item/8tp5x1hb (Accessed 15/06/21)
- Stewart, F. (2002). Root causes of violent conflict in developing countries. *BMJ*, [online], Volume 324(7333), p.342-5. Available at: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1122271/ (Accessed 06/03/21)
- United Nations (2001). *Report of the Panel of Experts on the Illegal Exploitation of Natural Resources and Other Forms of Wealth of DR Congo.* [online] Reliefweb. Available at: https://reliefweb.int/report/democratic-republic-congo/report-panel-experts-illegal-exploitation-natural-resources-and (Accessed 26/06/21)
- Uppsala Conflict Data Program (2020). *Armed Conflict by Type, 1946-2020.* [image] Available at: https://ucdp.uu.se/downloads/charts/graphs/png_21/armedconf_by_type.png (Accessed 03/07/21)

Electric Vehicles: The Political Cop-out

This article was published by the Financial Times on December 19, 2021

Cameron Allan, Upper Sixth

Most people believe the future of transport is electric vehicles. Their uptake has been increasing rapidly and governments are starting to create policies to encourage their use. But they are a cop-out which avoids difficult decisions by governments.

From a political standpoint, EVs are very easy to implement. They require no change in culture, comparatively little investment from central government and use existing road infrastructure. Most importantly, their adoption makes governments seem environmentally friendly.

With 27 per cent of UK greenhouse gas emissions originating from transport, there is a clear need to decarbonise. While aeroplane and shipping technology have not developed to the standard required, decarbonisation of cars can be achieved now.

But the problem is that EVs are not nearly as climate friendly as people believe. Some studies criticise the high level of emissions required to produce EV batteries and the electricity they use unless it is sustainably sourced.

Promoting car use creates other negative externalities such as increased congestion, which is estimated to have cost the UK economy £6.9bn in 2019. It reduces levels of physical activity and creates associated health problems, increases noise pollution and makes cities unpleasant to live in.

Decarbonisation will inevitably involve EVs, but walking, cycling and improved public transport should take centre stage. Keeping within the rules of the Paris climate agreement will require a change in culture, which is much harder to achieve than simply rolling out electric car subsidies.

It will require governments to restrict freedom, which is not easy in a democracy. The recent experience of Transport for London's Low Traffic Neighbourhood schemes (LTNs) is good example of the difficulties ahead. These schemes consist of traffic management schemes that close off roads to through traffic while allowing pedestrians and cyclists to pass.

They have been implemented to reduce "rat running" on local residential roads and push any remaining traffic on to main roads better adapted to cope safely with larger volumes of traffic. The result has been protests from people claiming there would be an increase in emergency service response times. Others have resorted to vandalism, with local groups damaging traffic counting equipment and even arson.

In many cases, the fear of not being able to drive everywhere evoked concerns about entrapment and exclusion from society. Other criticisms raised after the Sarah Everard case included perceived reduced public safety on blocked roads.

Yet one study found that the number of reported crimes in and around LTNs has shown changes "similar to or more favourable than the wider background trend". Waltham Forest, a borough with one of the most extensive schemes, identified dangerous levels of pollution fell and researchers from King's College London estimate life expectancy may have increased by up to six weeks compared to 2013 as a result.

A report by the London Fire Brigade found response times did not increase. Even fears of the main roads becoming permanently gridlocked have not materialised, with experts citing the "evaporating traffic theory" as drivers switched to walking or cycling. Levels of cycling in Waltham Forest have increased sharply.

Despite such positive findings, the borough was the scene of protests in 2015 when the schemes were announced. Unlike many councils which gave in to the loudest voices and watered-down their programmes, the council held its nerve. Now reportedly only 1.7 per cent of people want the road changes to be reversed.

Despite this plethora of evidence, newly installed LTNs are still often controversial, with one report joking that they are "more divisive than Brexit." Perhaps people are simply reluctant to change. The truth is that if we are going to solve the climate crisis, we cannot please everyone.

Using EVs as central to tackling climate change is a cop-out from governments wanting to avoid having to contend with difficult and sometimes unpopular decisions such as imposing LTNs and congestion charging.

Governments should focus their attention on discouraging car travel and promoting walking and cycling. In a country where people are currently used to driving practically everywhere, discouragement is by no means easy,

Recent data show that 27 per cent of UK greenhouse gas emissions come from transport, of which 61 per cent are from cars and taxis. With one-third of UK car journeys lasting less than five miles, discouraging car travel would go a long way to reducing the UK's carbon footprint — and far more quickly than EVs.

As we emerge from the coronavirus pandemic, scientists warn that we are entering another one: obesity. Active lifestyles including more active travel are key to the response. So why not kill tackle the climate crisis and obesity at once?

Walking and cycling is the transport of the future — but only if we are willing to stand up to resistance from people used to the freedom of driving where and when they please.

The RGS 4000 Characters Competition Winner

Joe Colton, Lower Sixth

STUDENTS IN THE LOWER SIXTH WERE ASKED TO EXPLAIN WHOSE WORKS THEY WOULD SAVE IN THE EVENT OF AN APOCALYPSE LEADING TO A COMPLETE RE-BUILD OF SOCIETY.

When considering the issue of the preservation of the works of a single author we have to be conscious of the purpose of this preservation. What are we aiming to preserve when we preserve this author's work? We should be aiming to preserve what gives people joy; maximising peoples' quality of life should be the fundamental aim of every world leader. So how best to maximise a future civilisation's quality of life? In accordance with the Easterlin paradox the most reliable indicator of happiness is income, a correlation which holds constant until very high-income levels. However, we cannot preserve income even as we know it today never mind what a future civilisation may perceive as income. It would also make no sense to attempt to preserve modern economic theory as it is rooted in the pre-conceived notions of states and industrialisation. Instead, what if we could preserve the values that underpin modern economic theory, which underpin our modern prosperous society, that we hold so dear. The values which are universally comprehensible. The values which we find in the work of Eleanor Roosevelt.

The pinnacle of Roosevelt's scholarship and the integral reason for why her work should be preserved is her role as the chief creator of the Universal Declaration of Human Rights. This is the comprehensive document for creating an open, tolerant society with the happiness and wellbeing of its citizens at the heart of it. The inalienable rights within it are set out in absolutes which can be easily grasped, and the rights are applicable within every society under every imaginable condition. This universality, the document being written without pre-conceptions or assumptions, is what propels Roosevelt's work above all others when considering what should be preserved. Key to her primacy is also the fact that her most important work, the UDHR, is clearly the preeminent one when compared to her other works. Its use of imperatives and formal layout distinguishes it from her other works. This is of paramount importance as with other scholars which work a future civilisation should pay particular attention to or disregard is not immediately obvious or not obvious at all. There is an ideological coherence throughout her work in favour of openness and tolerance and whilst it has been reported in her earlier years, she was a closet anti-Semitist not only did this not impact any of her works, but she also revoked this opinion well before she created any of her key pieces of scholarship.

To conclude, to preserve the values we hold dear in our open, equal, tolerant society; to preserve our modern age of prosperity we must preserve the work of Eleanor Roosevelt. For values which lay the framework for a better, kinder society without pre-conceptions or assumptions which can be applied to any society, past present or future, and transform it into a more just and equal society there is no scholar more effective than Eleanor Roosevelt. For values which are universally comprehensible and already accurately translated into over 370 languages and dialects there is no author on the same level as Eleanor Roosevelt. To preserve a more just, fair and equal world for the future preserve the work of Eleanor Roosevelt.

Discuss: "In the ancient world, the sea always linked rather than divided people."

Highly Commended entry to the Fitzwilliam College, Cambridge Classics essay prize

Mattie Sutton, Upper Sixth

The beginning of Herodotus' Histories neatly encapsulates the innate tension around the sea for people in the Ancient World: Phoenician traders arrive and 'set out their cargo'[1] - an exercise of economic benefit - but before they leave, 'Io and others' (local nobility) 'were seized and thrown into the ship'[2] - a politically incendiary device, which Herodotus tells us the Persians believe was the beginning of the East/West divide that would shape Greek history.[3] For the Ancient World, the sea was an area of opportunity, but also of threat - threat of piracy or invasion. The sea could link, but it could also divide.

This essay will examine possible links and divides in three key spheres - economic, political and cultural - taking our evidence from two areas: the Mediterranean and seas around Japan.[4] It is important, so we can draw broader conclusions about the impact of the sea in the Ancient World, to study two, isolated areas, and I decided to utilise my study of Classical and Japanese languages to allow for a comparison of the Mediterranean and the seas around Japan, a parallel which has scarcely been drawn previously.

Overall, economically the sea always links, however these economic incentives pushed states to more aggressive maritime policies, causing political divides. Finally, the sea and its economic interactions also provided cultural meetings, creating links, meaning the sea, while not always creating links between people, predominantly did.

Economically, the sea provided links around the Mediterranean, most notably in the Roman Empire. The fact that it was thirty times cheaper to transport by goods by sea than by cart pushed traders towards the more dangerous sea.[5] The Roman Empire then managed to harness this more convenient transport medium and allowed individual towns to specialise their output (e.g., wine but not grain). This enabled a move away from less effective local subsistence farming, giving way to a more linked trade system where necessary items were transported across the Mediterranean,[6] as evidenced by amphorae-carrying shipwrecks at Madrague de Giens,[7] Port Miou (which the Oxford Shipwrecks Database estimates could have over 1500 amphorae) and Le Grand Ribaud.[8] The geographical features of the Mediterranean allowed the people of the Roman Empire to conduct a more linked trading system.

In Japan, the economic incentives of the sea were creating similar links. Conditions caused by the warm Kuroshio current meant the Ryuku archipelago, south of the main islands, had unique access to tropical varieties of shells, used first for ornamental decoration but then became so important that they were statements of political authority. Archaeological evidence shows these shells ended up in the coastal and Yellow River regions of China, while other decorations, jewellery and coins from those regions have been found in the Ryuku archipelago, suggesting there was a comprehensive trading system.[9] As with the Roman Empire, where the efficiency of transport by sea enabled links by trade, the geographical feature of the warm sea current was empowering economic links by sea. The sea was not just the transport medium for the shells, but the source as well, similar to the Phoenician Murex trade,[10] which may have even developed a secondary industry in mainland Crete.[11] Furthermore in Japan, the introduction of agricultural techniques around 800 – 1000 AD suggest a further cultural and technological exchange between peoples, supporting Antonaccio's suggestion that links and networks carry not just 'artefacts, raw materials' but also 'ideologies, technologies'.[12]

However, the economic incentives of the sea quickly made it into another arena for political divides.[13] Rome was unique in its position of having the whole of the Mediterranean under its control and being able to call it simply '*nostrum mare*' ("Our sea") as Sallust and others refer to it.[14] At other times, as sea travel became a more enticing enterprise, political divides were caused by conflict and disagreements. Niemann's study of Shephelean coastline archaeology in the first half of the First Millennium BC has concluded that while their location

1. Hdt. 1.1.2
2. Hdt. 1.1.4
3. Hdt. 1.1.1, 1.2.1
4. It is worth noting Japan is not wholly united during our time period: one of our examples later will assess a divide between the islands that make up modern-day Japan, in addition to other examples assessing Japan's interaction with other civilizations.
5. (Buchet, Arnaud and de Souza 2017) P.620
6. (Culham 2017) P.284-5, (Buchet, Arnaud and de Souza 2017) P.622
7. (Parker 1992) P.249
8. (The Oxford Roman Economy Project 2010): http://oxrep.classics.ox.ac.uk/databases/shipwrecks_database/records/817/ http://oxrep.classics.ox.ac.uk/databases/shipwrecks_database/records/1460/
9. (Kinoshita 2003) P.67-72 https://journals.lib.washington.edu/index.php/BIPPA/article/view/11853/10480, (Buchet, Arnaud and de Souza 2017) P.108
10. The Phoenician's murex/purple trade is evidenced in *Plin. Nat. 5.17*, as well as *2 Chronicles 2 13-16* and secondary sources such as (Frangié-Joly 2016) P.51-53
11. (Carannante 2011) P.9
12. (Antonaccio 2017) P.217
13. (Buchet, Arnaud and de Souza 2017) P.624
14. Sal. Jug. 18, (Culham 2017) P.283

gave cities trading benefits, it also made them prime targets as wealthy and strategic positions,[15] meaning for any economic link a political divide may have been around the corner.

In Fifth Century BC Ancient Greece, an earlier, smaller example of the Roman economic interdependence also resulted in political divides. The Athenians had, according to de Souza, a 'long tradition of maritime trade'.[16] In his *History of the Peloponnesian War*, Thucydides reports that they were persuaded by Themistocles to finish Piraeus and develop a fleet, something Themistocles believes will provide a 'great advantage for the attainment of empire'.[17] Torelli among others suggests that this became a source of political divisions with other states.[18] The reason why it was an advantage was not only military, but because 'there is no city which does not have to import or export' as Pseudo-Xenephon (a 5th Century Athenian critic of democracy) states.[19] The Athenians as masters of the sea would therefore be provided with an economic stick with which to punish any dissidents.[20] Thus, the potential economic links of the sea could be weaponised and could create political divides between people, such as with the city of Aegina, who Strabo and Herodotus both report were long-time naval rivals of Athens,[21] and against whom the Athenians went to war around 460BC.[22] Athens' bold attitude when dealing with other states, enabled by the sea, made up a significant part of Thucydides' reported 'τας αἰτιας' ("the grounds of complaint")[23] which truly began the Peloponnesian War: a huge political divide. Additionally, the divisions caused by Athens' ambitions have parallels to the build-up to the First Punic War, when Rhegium helped draw Rome into intervention in Sicily by expressing fear 'of the Carthaginians as being masters of the sea', according to Polybius.[24] Ambitions for economic links could lead to political divides, and state's needs for these exports and imports could also be used against them, as demonstrated by Athens.

In 527 AD Japan, a rebellion in Kyushu, the most southernly of the main Japanese islands (and significantly the one nearest Korea), was similarly creating political divides by using the leverage that the Yamato court (the main power)'s need for economic links gave them. Staging a rebellion, our primary source the Nihon shoki specifically states how the leader Iwai (also called Ihawi) '勿使修職' ("did not allow the taxes [from his provinces] to be paid [to the Yamato]"), and '外邀海路。誘致高麗百濟新羅任那等國。年貢職船' ("Abroad he intercepted the route by sea, and led astray the yearly tribute ships from the countries of Koryo, Pekche, Silla, and Imna").[25] While the power dynamic is reversed from Ancient Greece, political divisions were created in the same way: by disrupting economic links. Commentators note that Iwai was rebelling in a particularly advantageous geographical position, giving them 'considerable control' over the trade route between Korea and the Yamato Court, meaning their disruption of economic links by sea could create political divides between people,[26] in a similar way to Ancient Athens.

The above paragraphs demonstrate the economic links and the political divides caused by the sea, but it is important to note that the sea could also link people culturally. A Roman example of this is the Lakshmi goddess statue found at Pompeii (pictured at end), which its home, the Museo Archeologico di Napoli, believes was of 'Indian production'. The Museum states that it 'represents an important indication of the trade relations... between the Western Mediterranean countries and the East', crucially 'by means of the port of Puteoli'.[27] The Lakshmi statue is an example of the religious cultural exchange that happened as a consequence of links between people that were originally forged as economic ones.[28] While not all cultural items would have been transferred by sea, its convenience as a transport medium combined with how these links were originally a product of Roman economic growth (also a result of the unique advantages of the Mediterranean) means we can safely suggest the sea was crucial in forging these cultural links.

Statue of the Hindu goddess Lakshmi found in Pompeii)

15. (Niemann 2013) P.245
16. (de Souza 2017) P.414
17. Thuc. 1.93
18. (Torelli 2003) P.120
19. Ps. Xen. Const. Ath. 2.3
20. (de Souza 2017) P.424
21. Strab. 8.6.16, Hdt. 5.82.1, and their naval interest is supported in Hdt. 5.83.1
22. (de Souza 2017) P.417
23. Thuc. 1.23
24. Plb. 1.7
25. *Nihon shoki* Bk. 17:18 (Keidai Tenno), English translation (P.15 of the Keidai Book): https://archive.org/details/nihongichronicl00astogoog/page/n452/mode/2up, Japanese (P.1015 of the Keidai Book on website): https://jhti.berkeley.edu/cgi-bin/jhti/select.cgi?honname=1
26. (Lee 2018) P.122, supported by (Hudson 2017) P.112
27. (Museo Archeologico di Napoli 2016): http://www.museoarcheologiconapoli.it/en/room-and-sections-of-the-exhibition/3490-2/
28. (Culham 2017) P.291 suggests interest in exotic products was the cause of the original interactions.

In Japan, similar religious cultural exchanges were happening as a result of economic activity at sea. The island of Okinoshima, north of Kyushu and south of Korea, was a religious centre where Japanese fishermen, sailors and traders would go to pray for safety at sea.[29] Archaeologists have found artefacts as well as 'exotic imported items' (some possibly acquired by maritime trade?) as offerings,[30] but importantly also Korean artefacts,[31] suggesting this was a religious centre for Japanese and Korean peoples where cultural links were created. Given the offerings were caused by a fear of the sea, and the geographical locations makes Okinoshima a meeting in the sea between Japan and Korea, it is clear these religious and cultural links were, similarly to the Roman ones, enabled by the sea.

Other cultural links between people were also facilitated by the sea, notably the development of languages. The Phoenicians, a Mediterranean trading powerhouse,[32] arguably made their greatest contribution by spreading their alphabet around with their trade (Paine refers to it as a 'high-value cargo' in itself), leading it to become the root of the Greek alphabet among others,[33] creating cultural links between people around the Mediterranean.[34] These cultural links were another consequence of the economic links created by Phoenician trading around the sea, based off the more convenient trading system that the Mediterranean allowed for.

In Fifth Century AD Japan, a similar linguistic transferral via sea was taking place. Writing was brought to Japan via the medium of Chinese books, brought over by a large body of people migrating from Korea.[35] Abulafia notes that 'turbulence' at home drove great numbers over, bringing with them writing.[36] While we think of the sea as a barrier in a negative sense, it allowed these people to properly escape the instability of Korea at the time, and consequently the sea forged a highly significant cultural link between Japan, Korea and China, in the form of a writing system.

To conclude, through our comparison of the independent Mediterranean Sea and the seas surrounding Japan, we have seen that economically the sea provides links between people, creating economic incentives such as efficiency and convenience that drove states to forge economic links. However, these economic links had consequences, and quickly made the sea into a political arena where the need for imports and exports could be weaponised by powers, either to subdue (Athens) or revolt (Iwai). Finally, economic links by sea also had the consequence of facilitating cultural exchanges by sea between civilizations that might not have otherwise happened: such as between India and the Romans, and Korea and Japan. The spread of ancient forms of writing, both not just by water but also incentivised by the sea (economically for Phoenicia, politically for Korea) further demonstrates cultural links that the sea enabled. The question proposed that the sea always linked rather than divided. However, in this essay I have shown that the sea did not always facilitate links between people in the Ancient World of the Mediterranean and Japan, particularly with regards to politics, but predominantly did, economically and culturally.

29. (Kaner 2012) P.55: http://www.okinoshima-heritage.jp/files/ReportDetail_81_file.pdf, (Abulafia 2020) P.191
30. (Kaner 2012) P.54-5
31. (Abulafia 2020) P.191
32. Isaiah 23:8 tells us Tyre, part of Phoenicia, had traders 'renowned in the earth'
33. (Cornell University Networks 2020), Herodotus notes the effect of the Phoenician alphabet on the Greek one in Hdt. 5.58.1
34. (Paine 2014) P.86
35. (Frellesvig 2010) P.11
36. (Abulafia 2020) P.192

The RGS Explore Competition 2021

Ollie Liversedge, Upper Sixth

IN THIS LINGUISTICS TASK, STUDENTS WERE ASKED TO TRANSLATE A PASSAGE INTO A LANGUAGE OF THEIR DESIGN

AKILLUMINAT

The 'Akilluminat' (Illuminati language) is the language of the Illuminati people. They are a reptilian race who disguise their behaviour and appearance in a humanoid form. They are able to imitate human sounds, but also have sounds which humans cannot make: the ∞ sound, which is a clicking sound made by constricting their windpipe; the ∫ sound, which is a high frequency snorting sound made by exhaling air through their papillae.

The Illuminati race evolved during a period in which the Roman Empire was at its greatest. As such, their language follows a grammatical system which is similar to Latin. However, as the Illuminati people do not want to be discovered, it is a mainly spoken language. When written down, the language does not use capital letters*; proper nouns are indicated by a line underneath.

Nouns do not have cases, the only time a noun changes is when it becomes plural: if the singular noun ends in a consonant, the plural noun will add the suffix '-i'; if the singular noun ends in a vowel, the plural noun will add the suffix '-de'.

*Numbers are recorded as Roman Numerals, and these are the only words written with capital letters in the language.

The written language uses punctuation: the beginning of a sentence is marked by |, a brief pause in a sentence is marked by \ (this replaces commas, brackets, colons, hyphens etc. etc.), the beginning and end of speech is marked by ¬.

There is no word 'to be', instead a compound is formed noun-adjective: if the adjective is always true for the noun, the compound is simple order noun-adjective; if the adjective is conditionally true for the noun (ie. not always true), the compound is joined by the letter ∞; if more than adjective describes a noun, the noun form an equal number of compounds to adjectives, separated by the letter ∫ (meaning and).

Verbs are only written in present tense. To conjugate the present tense, you take the infinitive and add '-n', '-ni', '-nin', '-d', '-di', '-din' (I, you singular, he/she/it, we, you plural, they) to the end. To indicate past tense, the word 'la' (before). To indicate the future tense, the word 'za' (after) is used.

PASSAGE 1 – ENGLISH: KURT VONNEGUT, GOD BLESS YOU, MR. ROSEWATER

Hello babies. Welcome to Earth. It's hot in the summer and cold in the winter. It's round and wet and crowded. On the outside, babies, you've got a hundred years here. There's only one rule that I know of, babies - "God damn it, you've got to be kind."

PASSAGE 1 - AKILLUMINAT

| hu wawade | é haka huhudi | í uppaha ret∞ha ∫ í uppaka ret∞ka | retwo ∫ ret∞sho ∫ ret∞hubbub | í haka \ wawade \ C azzi pa didi |I-I lex za \ ret baban \ wawade \ |¬ illuminat ret xa∫∞axni \ diora lexdi

DICTIONARY

('' indicates words translated from English)
hu = 'hello', greetings
wawa = baby
wawade = 'babies'
é = 'to', at, towards
í = 'in', 'on', into, onto
haka = 'Earth'
haka = 'outside', atmosphere, air
huhu = be welcome
huhudi = you are welcome, 'welcome (pl.)'
uppa = season
ha = 'hot'
ka = 'cold'
uppaha = 'summer'
uppaka = 'winter'
ret = 'it', thing, this, that, 'which'
reti = things, these, those, which
wo = 'round'
sho = 'wet'
hubbub = 'crowded'
∫ = 'and'
C = '100' (Roman Numerals)
azz = year
azzi = 'years'
pa = 'here', now
nene = I have
nini = you (s.) have
ninnin = he/she/it has
dede = we have
didi = 'you (pl.) have'
dindin = they have
za = after, 'there'
I = 1 (Roman Numerals)
I-I = 'only one', just one, one by one
lex = law, 'rule'
ba = know
baba = 'know of', understand, recognise, remember
illuminat = 'God'
illuminat = person, man, woman
xa∫∞ax = damn, curse
xa∫∞axni = you (s.) 'damn', curse
di = you (pl.)
ora = face, smile, 'kind'
lexi = to ought
lexidi = you (pl.) ought, must, 'have got to be'

Probing, Understanding and Optimizing Machine learning for Sparse data

Ishan Nathan, Upper Sixth

A SUPERVISED ORIGINAL RESEARCH PROJECT IN A SPIN-OUT COMPANY FROM THE THEORY OF CONDENSED MATTER GROUP AT THE UNIVERSITY OF CAMBRIDGE, INVESTIGATING A MACHINE LEARNING ALGORITHM THAT EXTRACTS HIDDEN FEATURES FROM SPARSE DATA IN MULTIPLE DIMENSION

ABSTRACT

In this paper, we attempt to understand, develop and optimize a machine learning architecture that extracts hidden features from sparse data in multiple dimensions. We analyse the current literature in machine learning and its inefficiency in dealing with sparse data. The result of which, we aim to optimize a certain hyperparameter to improve the R^2 value, in turn allowing the neural network-based model to formulate enhanced predictions.

The work presented in this paper is then applied to real-world experimental materials and pharmaceutical datasets in order to observe improvements in the accuracy of the regression model. We then discuss the benefits of the exploitation of sparse and noisy data in machine learning, as well how the improved extrapolation of target variables in sparse data can have wide-ranging applications within Physics and beyond, specifically financial markets.

1 INTRODUCTION

The discovery of materials has been part of quotidian life throughout human history; giving rise to the stone, bronze and iron ages, with each discovery enabling breakthrough technological advancements. However, limited by experimental conditions and theoretical foundations: experimental trial and error research consumes tremendous time and resources. Thus, the need for a novel method to accelerate the discovery and design for materials is required. A combination of both experiments and computer simulations are two conventional techniques that have been adopted into the field of materials science, yet it is difficult for these methods to accelerate material design and discovery.

Firstly, experimental trial and error measurements, although intuitive and easy to carry out, are conducted over a long period of time in an inefficient manner. Furthermore, computer simulations require high performance computer equipment and when a new system of data is studied, negligible use can be made from previous calculation results. In order to yield a strong understanding of the structures and properties of materials, machine learning tools have been adopted. This has enabled a substantial reduction in the time and cost for material design, producing vast amounts of data. Meticulously designing the right material for a specific purpose is essential. With the help of machine learning algorithms, handling typically sparse experimental datasets will become more time and cost efficient.

Machine learning is a branch of artificial intelligence (AI) and is solely focused on the use of data and algorithms to mimic the way that humans learn, gradually improving its accuracy. It is a method of data analysis that is reliant upon using existing data to uncover dependencies between various quantities. By finding patterns in high-dimensional data; it models the linear and non-linear relationships between the properties of materials, learning off previous empirical data. Moreover, it can understand the uncertainty in these predictions and can reveal new hypotheses for the conditions required for a successful product design/formation.

Notable examples of machine learning applications are molybdenum-base alloy design, nickel-base superalloy design, and alloy identification for direct laser decomposition; these designed materials outperformed all alloys that are commercially available and shows the power in optimizing bespoke materials for a given application.

Many real-world datasets represent domains where there are thousands of variables and multiple dimensions that are of interest. Traditional statistical analysis of such datasets typically require all or the vast majority of data points to be present. However, the vast majority of experimental data sets are sparse in nature, which means that there are many vacant data points in the data being recorded. One reason for their sparsity is partly because the information they represent is rare, but also because obtaining a full dataset of experimental results is very difficult to do so in laboratory conditions. This is because it is very expensive to measure and there are also different people with different motivations measuring various quantities.

Nonetheless, machine learning tools such as Alchemite™ possess a unique ability to handle data sets that are usually incomplete for input variables. Normally, these missing data entries would be discarded by deleting all the rows of data that do not have values, in turn making the data set complete. However, this method only works when one is training a machine learning model for one output column at a time. We, on the other hand, are insistent on training two or more output columns of sparse data simultaneously in order to exploit relationships between the two particular properties of the output columns. Therefore, we require the hyperparameter optimizer to select one of the hyperparameter values correctly and for the other column, we scale the value. The approach presented in this paper shows how the tool can exploit incomplete data and formulate enhanced predictions.

Within the pharmaceutical industry, sparse data AI is rapidly being adopted, to improve the predictive models in drug discovery. In this industry, however, much of the data generated is limited, resulting in a restricted amount of usable data. AI for sparse data, with detailed expert knowledge, augments experimental results for the probabilistic prediction of certain desired quantities. Other applications include neural networks capturing additional information, to be

used to design materials and to be used in 3D printing processes e.g., direct metal decomposition.

In this paper, we first look under the hood of the machine learning architecture, whilst investigating into a hyperparameter called *data for intercept* - validating the formalism in Section II. We are then well poised to delve into the importance of noise in Section III and develop this theory further in the analysis of sparsity in Section IV. Having derived a novel formula for analysing the relationship between the hyperparameter and sparsity, we apply this formula to real-life data sets within the materials and pharmaceutical industries in order to observe improvements to the accuracy of the regression model. After this, we further understand the algorithms potential application to other areas within Physics, in Section V.

2 MACHINE LEARNING

Machine learning algorithms formulate predictions of new systems to replace expensive and time-consuming experiments. They work by exploring data and identifying patterns, enabling the tool to formulate predictions of variables for new data entries. One way to predict the properties of a material, would be to fit a Taylor polynomial to existing datasets. However, being reliant upon polynomials, suffers from limitations such as oscillation tendencies and exploding values. Instead, Padé approximant are used; a rational function of numerator degree *m* and denominator degree *n*, where a_j and b_k are parameters.

$$y_{\text{padé}}(x) = \frac{\sum_{j=0}^{m} a_j x^j}{1+\sum_{k=1}^{n} b_k x^k}$$

[2.1]

The Padé approximant [2.1] converges very quickly and has an ability to find analytical structures within a function. This efficient tool to approximate complex real functions with only countable parameters, can be applied to many problems. By superposing Padé approximants, centred at different coordinates, one can fit a non-linear function to the data – known as a neural network. These neural network-based models can be applied to extract hidden features of sparse data in multiple dimensions. By using software like Alchemite™, which repeatedly fits curves to the data to obtain accurate models, we can train algorithms on fragmented data to identify missing correlations and make new material predictions. In this paper, we will not go into the details of Alchemite™, but we use the deep learning tool to create a formula that can optimize machine learning for sparse data.

The accuracy of these predictions are affected by certain hyperparameters, which are tuned, such as the amount of data used to calculate the intercept of a local straight line - data for intercept. The machine learning algorithm, tests different combinations of hyperparameters to train the optimum model for the dataset. In this paper, the datasets were trained using Tree-structured Parzen Estimator (TPE) hyperparameter optimization to identify the best hyperparameters. It follows a sequential model-based optimization approach (SMBO), which builds models based on historical data, to approximate the performance of a set of hyperparameters. It then subsequently chooses a new set of hyperparameters to test, based on this model. By finding the set of hyperparameters that minimise the mean square error in predictions on k-fold cross-validation, we can tune these hyperparameters. The mean squared error is the average sum of the squared difference between the actual value and the predicted value (12). A value close to zero will represent a greater quality regression model.

$$\text{MSE} = \frac{1}{n}\sum (y - \hat{y})^2$$

[2.2]

The R^2 value is the ratio of the Sum of Squares Regression (SSR) and Sum of Squares Total (SST) and its value measures how best the fit is. Hence a greater R^2 value, the better the regression model. In this paper, the focus is on improving the value to portray a more detailed picture of the regression model and to use this metric in assessing the predictive power of Alchemite™ on fewer data points. The calculation of the R^2 value is shown in [2.3]

$$R^2 = \frac{\text{SSP}}{\text{SST}} = \frac{\sum (\hat{y}_i - \bar{y})^2}{\sum (y_i - \bar{y})^2}$$

[2.3]

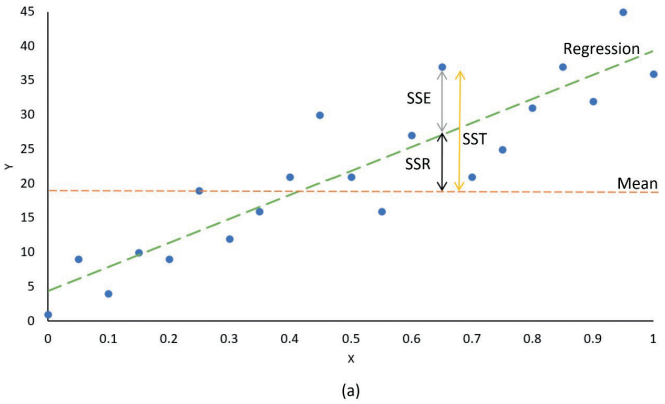


Figure 2: (a) shows the computation of the R^2 value. The green dotted line is the regression best fit line, the orange line indicates the mean of the actual/response variable value. The grey line shows the sum of squares residual error (SSE), the black line shows the sum of squares regression (SSR), and the yellow line denotes the sum of squares total (SST).

In order to estimate the skill of machine learning models on unseen data, the k-fold cross-validation procedure is adopted. In this cross-validation method, the training data is arbitrarily split into k subsections with different 80/20 splits (where a random 80% of the data is used to train the model and the remaining 20% is used as the validation dataset to test the model trained on the remaining subsections). A mean is taken of the k repeated results, giving the average R^2 value (which ranges from -∞ for randomly inaccurate predictions to 1 for ideal predictions).

When analysing experimental data, one has to take into account noise - a random bit of disorder that is present everywhere. It is a random error, or statistical uncertainty arising from a range of factors associated with the experiment one undertakes. For example, when plotting a meter reading for the output voltage from an ohmmeter, there are fluctuations which arise from the thermal motion of individual atoms. Or when conducting a radioactivity experiment, the intensity is monitored by a Geiger counter, here the statistical uncertainty emerges from the radioactive emission of particles due to its quantum-mechanical uncertainty. Noise is of paramount importance and is an essential ingredient of our physical universe, which needs to be understood and appreciated. For a given data set, to model noise on paradigmatic data sets, we firstly generate multiple machine learning models and then take the average of them to obtain the expected value. We then take the standard deviation of them to get an uncertainty. The noise can then be used to calculate the likelihood that a material will satisfy the criteria – an important feature to consider for materials design.

Machine learning tools such as Alchemite™ possess a unique ability to handle data sets that are usually incomplete for input variables. They find correlations between properties to help train a better model. They are also proficient in training multiple machine learning models. They do this by taking the initial training data set and withdraw from it a new training data set which has the same number of entries. It draws points at random, from the original dataset and with replacement. This implies that some data points may appear twice or thrice and some may appear zero times. Producing new training datasets, generated from sampling the original data set uniformly with replacement, is a common method in machine learning – known as bootstrapping. The Figure 2 (b) shows an example of how the machine learning algorithm works in predicting a value of Y from a given value of X.

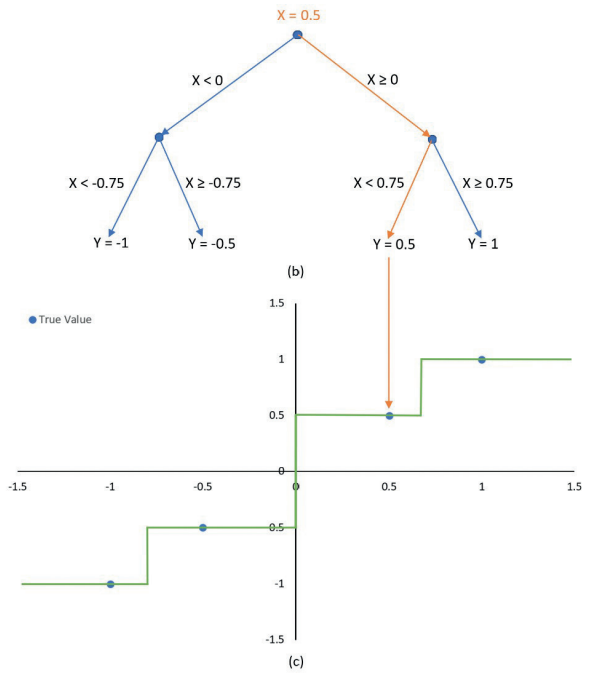


Figure 2: (b) outlines an optimal binary tree for predicting Y given X and (c) shows the plot in which the blue points are training data and the green line is the prediction.

3 NOISE AND UNCERTAINTY

Now that we are well versed with the machine learning methodology, we are well positioned to go and investigate fitting machine learning to noisy data. Noise is a key component of this paper, and we need to confirm whether our neural-network is robust against noise before we build upon this knowledge and optimize the machine learning algorithm for sparse data.

In order to confirm that the algorithm gives a good uncertainty estimate, we created a dataset comprising of N=100 inputs where $0 < x < 100$ is normally distributed $X \sim N(\mu\sigma^2)$ with $\mu = 0$ and $\sigma^2 = 1$. Below is the fit to this data set.

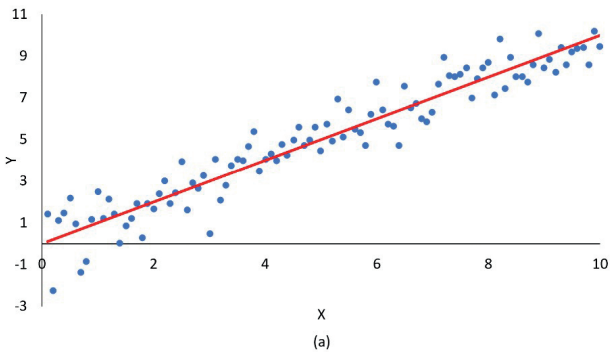


Figure 3: (a) shows the linear function of equation y = x , with Gaussian noise σ = 1.

As shown in Figure 2 (c), we can see that the predictive model, relies upon making a series of steps. In order to perform this, and before any consideration of sparsity, we need to decide what the averaging length-scale of those steps is, and we have to assess what the effects would be if we changed the length of these steps. For analysis of the averaging length-scale we use a data set with noise 5 and consider $Y \sim N(X, \sigma^2)$, assuming the optimum *data for intercept value is n.*

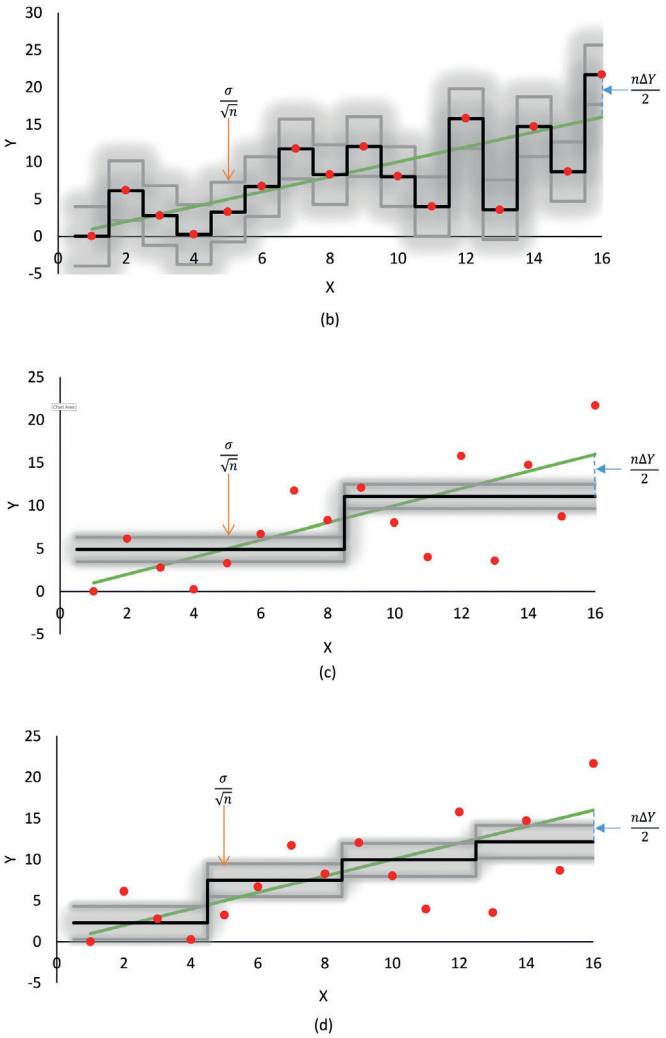


Figure 3: Graphs portraying systematic shift (b) small average length-scale steps, whilst (c) shows larger steps. (d) shows the original data set with the perfectly scaled steps (in black). All data sets have the actual values in red and the trendline for the data set is denoted in green. The grey shaded area is the mean square error.

From Figure 3 (b) we can see that as we make the scale of the steps too small, the data is scattered leading to the algorithm averaging a curve anywhere over the data. Here, we can see that the black line is being pulled away from the green line by the noise. The typical distance that it is being pulled away by is the contribution of the mean square error from white noise denoted by $\frac{\sigma}{\sqrt{n}}$ (the grey shaded region of the graphs). This value is the standard deviation of the points it encompasses hence the error in the mean. In Figure 3 (b) this value is too large as the black curve heavily fluctuates.

On the other hand, if the scale of these steps is too large, shown in Figure 3 (c), we are then averaging over the whole data set. Here, we can observe that the distance between the green line and the black line is very large. This distance, $\frac{n\Delta Y}{2}$, is our value for the average contribution to the cross validation mean square error, where $n\Delta Y$ is the increase in the linear function (the total amount the green line goes up across the horizontal plateau). The benefits of this model would be that there would be little/no uncertainty. However, it does imply that

by averaging over too many points, the function would have changed a lot over that scale and thus we have another inadequate approximation.

Therefore, in order to fit a good model, we need to find the ideal balance between uncertainty and systematic shift – shown in Figure 3 (d). To do this mathematically and obtain the total squared error: we add the values in quadrature to obtain the net variance in [3.1].

$$\left(\frac{n\Delta Y}{2}\right)^2 + \frac{\sigma^2}{n}$$

[3.1]

Taking the derivative with respect to n and minimising the expression we end up with:

$$\frac{\sigma}{\Delta Y n^{\frac{3}{2}}} = \frac{1}{\sqrt{2}}$$

[3.2]

Using a dataset of *N*=100 inputs and experimenting with different values of noise (σ), finding the optimum value of *n* using the *k*-fold cross validation method with *k=5*, we found the mean value of $\frac{\sigma}{\Delta Y n^{\frac{3}{2}}}$ to be 0.699 (which lies within one standard deviation of $\frac{\sigma}{\Delta Y n^{\frac{3}{2}}}$ from $\frac{1}{\sqrt{2}}$), shown in Figure 3 (e).

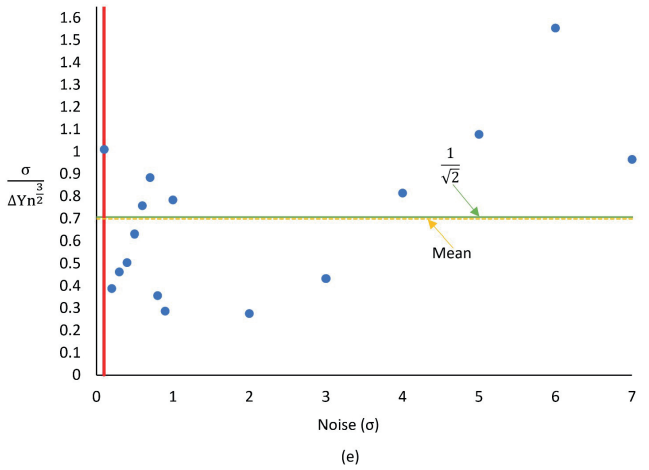


Figure 3: (e) shows different values of noise in blue, the actual mean in yellow and the theoretical mean in green. The red line indicates the lower bound for the noise level.

Further to this, we decided to research into the hyperparameter *data for intercept* which is the amount of data used to calculate intercept of local straight line. We used this hyperparameter as opposed to others as they did not target noise averaging. In order to investigate noise/uncertainty further, we manually changed the value for *data for intercept* to see how the quality (R^2) varied. The maximum on the graph indicates the ideal value for *data for intercept*.

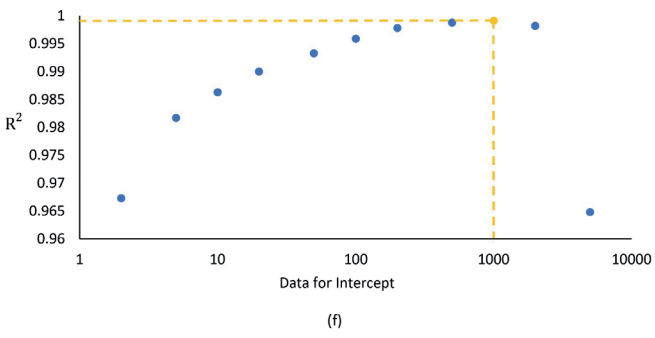


Figure 3: (f) shows how quality of the data varies with data for intercept. The yellow point indicates the optimum value for data for intercept.

4 SPARSITY

The vast majority of experimental data sets are sparse, which implies that there are many empty data points in the data being recorded. Normally, these data entries would be discarded, but having validated the machine learning tool for noise/uncertainty, we have a solid platform to do the same for sparse data and confirm the capability of the multilayer regressor, which utilises target variables for extrapolation. Sparsity is homogenous to noise in that normal experimental data is sparse. This means that there is more information clustered in one place in comparison to another, we are able to compute uncertainty from this absence of data.

Using the same method as we previously did in Section 3 to obtain the total square error, we add the two quantities in quadrature again in [4.1]. However, this time we need to consider the sparsity of the dataset, i.e s = 0 (all points are present) and s = 1 (there are no data points). We plot the same graph as we did previously in Figure 3 (d), except half of the data points are now missing. Whilst this has no effect on the error of the function, the error in the noise scales as $\frac{1}{\sqrt{1-s}}$. Therefore, we now have more noise as there are fewer points that we are averaging over.

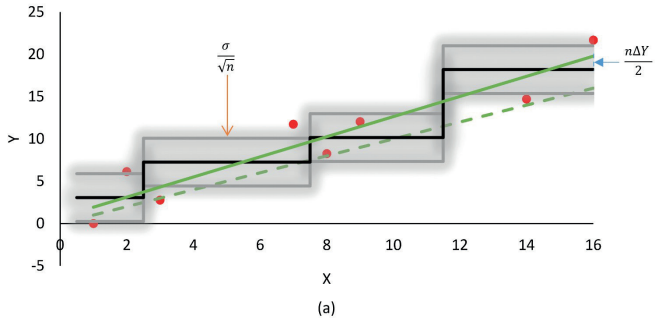


Figure 4: (a) displays the real sparse data set with perfectly scaled steps (in black). All data sets have the actual values used in Section 3 in red; the error shaded in grey; the original linear trend line with all the data points (dotted green) and the new trendline for the sparse data set in green.

As seen in Figure 4 (a) the uncertainty error bars $\frac{\sigma}{\sqrt{n}}$, denoted by the shaded grey area, differ over each horizontal plateau. This is because, the number of data points (*n*) that the line is averaging over changes and therefore $\frac{\sigma}{\sqrt{n}}$ changes too. Over the whole dataset, the average error is out by the sparsity constant $\frac{1}{\sqrt{1-s}}$.

As we did in Section 3, we add the two quantities in quadrature whilst this time considering sparsity too.

ΔY is the ratio of the range of data over the total number of data points within the dataset; σ is the noise level and n is the hyperparameter that we are interested in (data for intercept).

$$\left(\frac{n\Delta Y}{2}\right)^2 + \left(\frac{\sigma}{\sqrt{n(1-s)}}\right)^2$$

[4.1]

Taking the derivative, minimising and solving the expression with respect to n we get:

$$n = \left(\frac{2\sigma^2}{\Delta Y^2 (1-s)}\right)^{\frac{1}{3}}$$

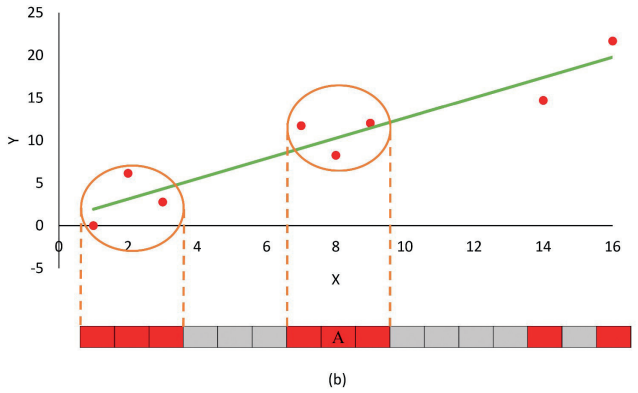
[4.2]

We tested this equation on the relationship between the hyperparameter of interest (*data for intercept*) vs sparsity. We then calculated the inverse of the normal cumulative distribution for the specified mean and standard deviation with a randomly generated probability, mean of zero and with the standard deviation being the noise in this prediction (5 in this case). Having plotted the values of *data for intercept*, it appeared that the formula in [4.2] mapped the values well for low levels of sparsity, but not higher levels – the mapping is shown by the blue curve in Figure 4 (e-f).

We are now going to look into the additional consequences of sparsity and reasons for the inaccuracy of formula [4.2] when modelling high levels of sparsity. Below we analyse two of these consequences. The first reason for this inaccuracy is because there are not enough points left in the dataset, which means that we cannot fit an accurate approximation.

The second consequence is that some of the points are not in a uniform distribution. Instead, they appear to be in a length-scale that poses some significance. As seen in Figure 4 (b) the noise and presence of the points are becoming more clustered due to sparsity. If the *data for intercept* goes from one of those points, over the large gap, to the next adjacent point, it is travelling a long distance and hence the function is susceptible to change over this length-scale. This change is much greater than we expected in our previous analysis and so we have to consider an additional constraint for this emergent length-scale. This is important as we do not want to fit over more data than available as we will be fitting over empty space, and we do not want to be fitting over missing data as the data will be too noisy.

Therefore, in order to try to fit over the clusters and make a comparison between the change in function over noise we revisit Figure 4 (a) and analyse it more carefully than we did before. We then begin to notice that we do not only have less data, but it appears the data is in groups – introducing an emergent length-scale. This is shown in Figure 4 (b).



(b)

Figure 4: (b) displays the emergent length-scale and clusters in the sparse dataset.

Expanding on this, to find a mathematical relationship of this length-scale, we need to work out the expected value/ length of these clusters. We begin by choosing an arbitrary valid starting point within a cluster in the data set, denoted by, "A". We then work out the probability distribution to the right of this point – given by $(1-s)^n s$. If we sum this over n, it equates to 1 as all probabilities add up to 1. To then obtain an expectation value for this probability distribution, we need to multiply this value by n to acquire $\sum_{n=0}^{\infty} n(1-s)^n s$.

Thus far we have worked out the probability to the right of point "A", but we also need to consider a separate probability distribution going to the left of this point. We multiply the sum above by 2 to compensate for this, and we then add 1 to this to allow for our known starting value "A". Combining the above, we derive equation [4.3] and hence [4.4], which can mathematically describe the clusters length-scale in Figure 4 (b).

$$1+2\sum_{n=0}^{\infty} n(1-s)^n s$$

[4.3]

Which simplifies to:

$$\frac{2}{s}-1$$

[4.4]

We then use equation [4.4] and regularize it with [4.6] to create our equation for a more accurate model for higher levels of sparsity shown in [4.7]. This equation is mapped by the yellow curve in Figures 4 (c-d).

$$\frac{\sigma}{\sqrt{n}} = n\Delta Y$$

[4.5]

Rearranging for n we get:

$$n = \left(\frac{\sigma}{\Delta Y}\right)^{\frac{2}{3}}$$

[4.6]

$$f(s) = \left(\frac{2}{s}-1\right)\left(\frac{\sigma}{\Delta Y}\right)^{\frac{2}{3}}$$

[4.7]

After obtaining different functions described by [4.2] and [4.7] we need to somehow combine these two equations to get a single function that can map our *data for intercept* vs sparsity relationship.

Let's recall function [4.2], calling it $f(x)$

$$n = \left(\frac{2\sigma^2}{\Delta Y^2 (1-s)}\right)^{\frac{1}{3}}$$

and Function [4.7], calling it $g(x)$.

$$f(s) = \left(\frac{2}{s}-1\right)\left(\frac{\sigma}{\Delta Y}\right)^{\frac{2}{3}}$$

In order to combine these two equations, we need to satisfy the following relationships.

$$y \to f(x) \quad if \quad f(x) \to \infty$$
$$y \to f(x) \quad if \quad g(x) \to \infty$$
$$y \to f(x) \text{ or } g(x) \quad if \quad f(x) = g(x)$$

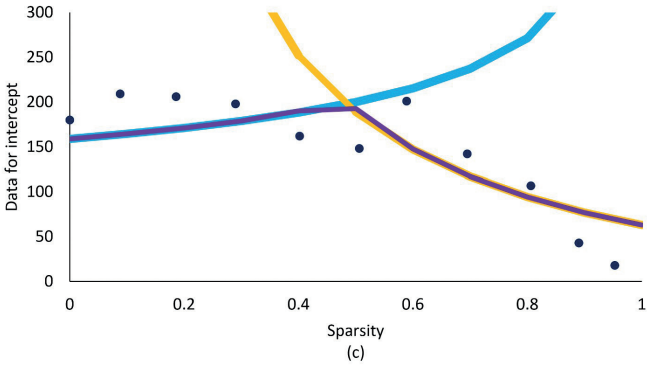
[4.8]

Firstly, in order to do this, we considered an appropriate way to combine two functions and find the average rate. We propose a function that encompasses the properties mentioned in [4.8], but also has practical purposes, enabling us to code the formula. By considering another form of average and rearranging, we obtain equation [4.9]. This formula combines both functions f(x) and g(x) very well and allows us to accurately map the relationship in Figures 4 (c-d) shown by the purple function. When considering [4.9] an obvious limit is the

lower bound of two. This is the pre-determined minimum value that *data for intercept* can be. This is because it has to be at least the number of data points in our entire dataset, otherwise, there is not enough information to fit a curve. In the equation [4.9] we use z an arbitrary scaling power to produce an accurate fit. The typical scale is 0 < z < 15 so that the function swaps between the two functionalities on a similar length-scale to that of the levels of sparsity.

$$\frac{f(x)g(x)[f^{z-1}(x)+g^{z-1}(x)]}{f^{z}(x)+g^{z}(x)}$$

[4.9]



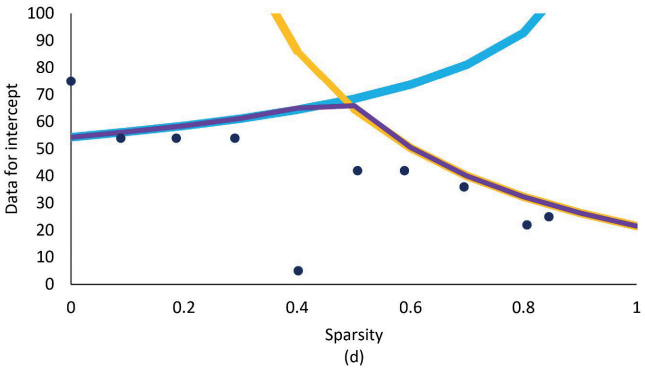


Figure 4: Graph to show the relationships between data for intercept vs sparsity for different size data sets and different levels of noise (. (c) Linear data set with N=10,000 entries, noise 50 and varied values of sparsity. (d) Linear data set with N=1,000 entries, noise 10 and varied values of sparsity.

As shown above, the purple function in Figures 4 (c-d) accurately maps the relationship between *data for intercept* vs sparsity for higher number of inputs and varied levels of noise. We used a noise level of 50 for 10,000 data points (scaled from 0.5 for 100 data points) and used a noise level of 10 for 1,000 data points (scaled from 1 for 100 data points), both of which are in the range shown in Figure 3 (e). By looking at the ratio in the function in [4.9] it is evident that it is independent of $\frac{\sigma}{\Delta Y}$ as this value cancels - concluding that the function will always work as it only relies upon the sparsity of the data (which can be measured prior to testing).

To prove that this formula can be applied to the machine learning algorithm's source code, we performed additional testing upon non-linear functions and data sets with higher dimensions: shown in Figure 4 (f) and Figure 4 (g) respectively.

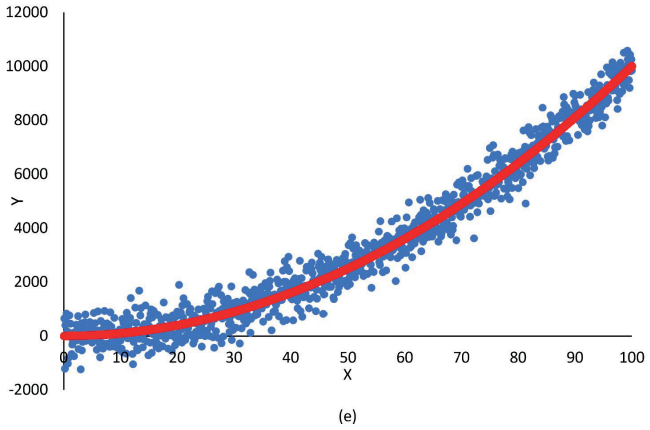


Figure 4: (e) displays the parabolic function with Gaussian noise σ=500 (blue) and the actual function $y=x^2$ plotted in red.

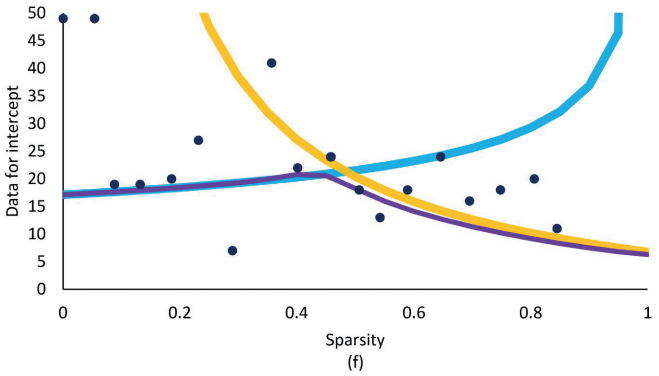


Figure 4: Graph to show the relationships between data for intercept vs sparsity for different size data sets and different levels of noise (σ). (f) Parabolic data set with N=1,000 entries, noise 500 and varied values of sparsity.

We also tested the algorithm on a hyperplane formed by using three random variables (x_1, x_2 and x_3) and summing them together to obtain our function.

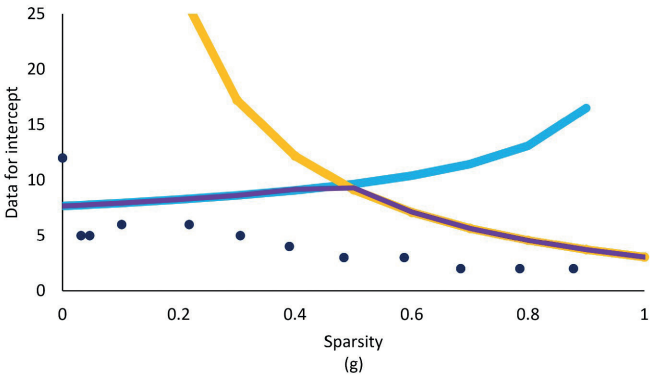


Figure 4: (g) Hyperplane with three input variables. N=1,000 entries, noise 0.15 (5% of scaled range of data) and varied values of sparsity.

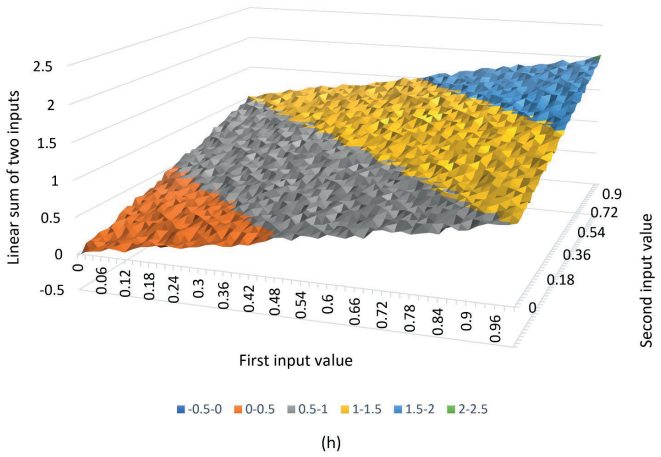


Figure 4: (h) Shows a pictorial representation of a hyperplane plot with two input variables.

5 APPLICATIONS TO PHYSICS AND BEYOND

Having set up an equation that accurately maps the relationship between *data for intercept* and sparsity, improving the R^2 value. We can now portray a better picture of the regression model and use this metric in assessing the predictive power of the machine learning tool on real-life materials and other applications in Physics. We test the algorithm on two different systems: a linear paradigmatic data set and a commercial dataset from a leading pharmaceutical company. We then further discuss how we can possibly use experimental data from the literature to test the formalism provided we had access to richer sets of real-world data.

Throughout the paper, all the datasets we have tested on only include single output columns. So, we decided to test the algorithm on a paradigmatic data set of over 10,000 rows of data and with two output columns. The initial R^2 value was 0.9036 but after the algorithm was implemented the new R^2 value was 0.9130. Normally, it is very difficult to render such improvements in the accuracy of this data, yet this algorithm demonstrated a 10% increase in the accuracy.

Further to this, we tested it upon several real-world pharmaceutical and material data sets. In all models, there was either an improvement in the R^2 value or it returned the exact same value as before, meaning that it did not weaken the value in any way. One notable example of it improving the R^2 value was on a data set from a leading pharmaceutical company. Historically, the R^2 value for this dataset was 0.7528, but after implementation it increased to 0.7580, indicating an improvement in the accuracy. This dataset had over 20 inputs and over 150 output columns with around 20,000 rows of data, the majority of which were sparse. The novel formula in this paper has now been pushed through into the production version of the Alchemite™ neural network source code.

Having seen the successes of the algorithm in the pharmaceutical and materials industry, we are now well-positioned to investigate further into the applications of machine learning for sparse data within other branches of Physics. One such field is the discovery of dynamical systems from experimental data as this would intrinsically enhance generalisation and extrapolation. By computing an algorithm that is capable of identifying the links between variables through physical relationships, it would aid extrapolating the system behaviour beyond the training domain. Many fields within Physical science would benefit greatly from this property: such include the aerospace industry, since many problems arise due to component failures and other related anomalies reducing the system performance. By analysing the limited real data that is available for training and validation, we can adopt the machine learning algorithm to create powerful and interpretable models.

It would also be interesting to apply the method to richer sets of data within the field of materials science. In particular, applying this method to experimental material datasets which are inherently sparse in nature, to quantitatively predict certain variables that may have been missed by conventional machine learning methods. One other area that the algorithm has potential applications in is modelling financial markets. Here, we can use the machine learning architecture to analyse the uncertainty of stock price fluctuations where the data is relatively sparse.

6 CONCLUSION

By understanding, optimizing and validating a machine learning algorithm specifically for sparse data, we were able to create a novel formula that can be adopted into the machine learning source code. In turn, allowing the neural network-based model to extract hidden features from sparse data in multiple dimensions; identify missing correlations; improve the R^2 value and make enhanced model predictions.

This algorithm was tested on multiple paradigmatic datasets, from simple linear models with varied levels of noise (σ) to non-linear functions, as well as those with a greater number of inputs/outputs and higher dimensions. The fact that σ and Δy have the same units means that ratio $\frac{\sigma}{\Delta Y}$ cancels out dimensionally and hence the formula only relies upon the sparsity of the data – a variable that can be measured before predictive model testing. Further to this, in this particular formula, we are fortunate that within the function the quantitative value for $\frac{\sigma}{\Delta Y}$ cancels out with the same quantity in a different output column. This concludes that the equation will always work, even if we do not know what the ratio between the two values are.

The hyperparameter optimizer selects *data for intercept* for the column with the most data in, and for all the other output columns, it will scale *data for intercept* according to the level of sparsity. Prior to the derivation of this formula, Alchemite™ assumed that the values for both $\frac{\sigma}{\Delta Y}$ and sparsity (s) were the same for all output columns. However, by implementing an algorithm that can accurately map the relationship between *data for intercept* and sparsity, the model is now able to make improved predictions and now only assumes that the ratio of $\frac{\sigma}{\Delta Y}$ is the same. In comparison to what Alchemite™ previously did, this is an improvement and lays the template to explore the relationship when the ratio of is not the same between output columns. We then can also go beyond this to optimize other hyperparameters in order to improve the R^2 value further. This endorses the method's applicability to more complex systems such as attempting to optimize skewness and kurtosis, which are higher order parameters that provide information on the shape of a probability distribution.

In this paper we have seen that the algorithm has shown improvements to the regression model in multiple real-world pharmaceutical and materials data sets. It also has the potential to be applied to richer experimental sets of data and other machine learning methodologies, beyond the physical sciences and financial markets.

Nudges – Marketing or Mind Control?

Tom Wright, Upper Sixth

INTRODUCTION

Nudges are apparent throughout all aspects of life, acting as a hidden force and influencing a significant part of how we live our lives and make decisions every day; from which drink you choose to the way in which governments ensure people socially distance during a pandemic. Put simply, a nudge can be viewed as small deliberate actions made to intentionally influence an individual's decision-making process. An example of a nudge could be in the layout of a supermarket, depending on where items are placed in the shop changes people's perspective on them. For example, people will subconsciously associate the most prominent products as being 'better' when in reality there is no rational or logical reasoning for this, this assumption is known as a heuristic or a 'mental shortcut.' Heuristics are essential as they ease the cognitive load of making a decision. A nudge works by utilising these heuristics to lead a consumer to follow a desired outcome of a 'choice architect'. The choice architect is the individual who designs the nudge and the conditions under which an individual would make their decision. Companies who utilize nudge techniques to improve the sales of their products or services result in a 14% increase in customer retention. (Anna Güntner, 2019) It is clear nudges are a powerful and lucrative tool when used effectively, so it begs the question are nudges just an effective marketing technique or are they form of mind control.

To determine whether a nudge can be classified as a marketing technique or a form of mind control, we must first understand what key aspects define a nudge and then analyse to what extent this meets the criteria of a marketing technique or that of mind control. Richard Thaler, often seen as the founder of behavioural economics (The Big Short, 2016), alongside Cass Sunstein stated the definition of a nudge to be "any aspect of the choice architecture that alters people's behaviour in a predictable way without forbidding any options or significantly changing their economic incentives" (Thaler, 2009). There are two fundamental aspects here which define a nudge: 1) that it "alters people's behaviour", 2) "without forbidding any options". In order to identify whether a nudge can be considered either marketing or mind control, we need to understand exactly what aspects characterise each. According to the American Marketing Association, marketing can be defined as "the activity of set institutions and processes for creating, communicating, delivering and exchanging offerings that have value for customers, clients, partners and society at large" (AMA - Association, 2017) This definition infers that the key principles of marketing are that it is only an offer and must be beneficial to the customers. In contrast, mind control can be defined as "systematically using unethically manipulative methods to persuade others to conform to the wishes of the manipulator" (Definitions. net, 2020). We can deduce the key principles that underpin mind control are its unethicality and manipulative nature. In order to determine which classification a nudge more closely aligns to; I will evaluate to what extent the principles of a nudge are shared with that of marketing or mind control.

OFFERINGS

The first key principle of marketing is that it must be an 'offering,' so as long as the recipient ultimately reserves the right to choose either to accept or to reject what a company is offering, then a nudge will more closely fit the definition of marketing. By its nature, a nudge must operate "without forbidding any options" (Thaler, 2009). In contrast mind control would suggest that the recipient of the nudge would be forced into making a certain decision. Take the status quo bias for example (Muthoo, 2020), such as in the UK pension scheme. Creating a pension is pre-selected, so an individual must actively decide against this if they would not like to make that decision, on an opt-out basis. This is classified as a nudge as it influences people's decisions while not preventing either option. It would suggest that a nudge could be seen as marketing as, whilst it encourages a certain action, it is still an offer, and the recipient has not been forced into a predetermined decision.

In this example, if a nudge were to be seen as mind control, then the individual would be unable to opt-out and would be forced to join the pension scheme with no alternative. It would also directly conflict the philosophy of libertarian paternalism, which all nudges stem from. Formed by Thaler and Sunstein, the movement acts to prevent the negative growth of nudges into mind control, incorporating the two key aspects of a nudge. Firstly 'Libertarian,' by always ensuring that the recipient never has their freedom of choice compromised (Harvard Library, 2014) and secondly 'Paternalism' meaning that the nudge must be beneficial to the recipient as they see fit. As this philosophy was formed by the same individuals responsible for the popularisation of nudges, their opinions can be deemed credible due to their deep understanding of the topic. To sacrifice one of these principles would mean that the action could no longer be considered a nudge. As long as a nudge upholds these principles then a nudge can only be deemed as a form of marketing.

BENEFICIAL

The second key principle of marketing is that it must 'offer' something of value and to be 'beneficial' for the customer, which a nudge does. No matter what the circumstances are, the offer which is beneficial to the customer remains throughout. The customer will still be faced with the same decision whether there is a nudge with a motive behind it or not. A company can present the offer to the customer in whichever way they see fit, so long as it complies with the law, it becomes the customers responsibility to make a decision whether to take up the offer or not. (Alemanno, 2014) Although nudges do influence decisions that the customers make the trade in itself is unchanged by a nudge and remains a beneficial offer and therefore continues to align more with the qualities of marketing as opposed to that of mind control which by its nature would likely be unfavourable to the customer.

In addition, companies are unable to remain neutral when making decisions that will affect customers, can they really be held accountable for prioritising personal gain over customer benefit? It becomes impossible for a choice architect within a company to remain completely neutral, so in many cases companies will nudge customers irrespective of their motives. Continuing from the above example of nudge in the layout of a supermarket, imagine you are the choice architect and you are faced with the decision of whether to place the unhealthy drinks at the front which are more popular and will result in a greater

number of sales, or placing the healthy drinks there promoting a healthier lifestyle to many customers. No matter what option you choose, one way or another you will be nudging the customers. It would seem if you were to randomize which drinks were shown first this would be completely neutral (Thaler, 2009). In one sense this is a neutral position but in terms of nudges it is not as you have made the decision to randomize which has resulted in some customers being nudged towards the healthy option in some stores and towards the unhealthy option in other stores, you are therefore not neutral as you are still nudging people in certain ways no matter if the motive was to be impartial. As it is impossible not to nudge then can companies realistically be called unethical for promoting sales that benefit them as opposed to choosing not to at a detriment to themselves. (Jespersen, 2017) By the fact that a nudge incorporates the beneficial trade aspects of marketing within it and that companies are unable not to nudge, it demonstrates that a nudge can only be classified as marketing.

UNETHICAL

The qualities which underpin how a nudge works are intrinsically unethical, which would suggest that it is a form of mind control. Latest studies indicate that over 95% of decision making takes place in the subconscious mind. (Zaltman, 2003) This statistic is not surprising as so many of the repetitive tasks done every day are achieved while not needing to engage our conscious brain. For example, commuting to work or school begins to become 'second nature' or a reflex action and no longer requires a conscious input to guide decisions. This is useful but is also an inherent flaw within the human cognitive process. Nudges can only work when the decision is made subconsciously (Jespersen, 2017) where aspects of rationality are lost, and the brain is guided by heuristics over rational and logical thinking. It can then be deemed unethical to target this vulnerability in the operation of the human brain, utilizing an individual's own heuristics to further the economic growth of companies at a detriment to the consumer, demonstrating that nudging is morally wrong and unethical (Ewert, 2019).

The most significant investigations into mind control were conducted by the US during the cold war, which has widely been highlighted as unethical. Under the MK-Ultra top-secret program the US covertly tested 'special interrogation' methods often involving high amounts of LSD to effectively force individuals to reveal secrets (Kinzer, 2019). Many individuals had their minds altered without giving permission, although there is no official documentation of this, it is believed that many were psychologically tortured and died as a result. In terms of the ethics between these actions and nudges, there are many key principles that overlap such as the fact that people had their behaviour altered and that they did not give permission for this to happen. However, the application of these principles is vastly different, and I am not suggesting that when companies nudge customers this has the same impact as the experiments that the US conducted, it definitely does not. Instead, it highlights that there are clear links behind the principles that both are based on, utilizing the vulnerability within the human cognitive process to control the decision-making process. (Campaign, 2004) Although it is true that companies cannot be deemed unethical for nudging based on the fact that it is impossible not to, it is also evident that the nudge, through the mechanism by which it works, is still inherently unethical and demonstrates it to be a form of mind control.

MANIPULATIVE

The final characteristic of a nudge by which it can be seen to resemble mind control is the inherently manipulative nature of not informing the recipient that their behaviour is being changed. The lack of transparency that underpins the key principle of a nudge can not only be seen as unethical but also deliberately manipulative, which are also common traits of mind control. (Engelen, 2015). An extreme version of a nudge could include the use of subliminal advertising which involves flashing images at recipients with the aim of making them do or think whatever has been suggested. Since 1958 this has been made illegal under UK legislation (Charles, 1958) "No advertisement may use images of very brief duration, or any other technique which is likely to influence viewers, without their being fully aware of what has been done". (BCAP, 2002) However, despite being a law, studies have proven that it is ineffective at making individuals follow the desired outcome (BBC, 2015). At the height of the cold war and the rise in fear of psychological warfare it is more than likely that this acted simply to prevent social unrest rather than being a legitimate threat. It does indicate that individuals being unaware of their behaviour being altered was unacceptable and was seen as unjustly manipulative, because individuals were afraid of being controlled without knowing it. The deceptive nature of this nudge resembles that of mind control due to the hidden motives behind both.

In order to demonstrate the power of manipulation, an example popularized by Dan Ariely, highlights the pricing strategy techniques companies employ. A news company offers three purchasing options to read its articles: Internet subscription (£50), a paper subscription (£110), a paper and internet subscription (£110), all prices are yearly contracts. To most individuals this pricing strategy would seem odd because it means no one would buy a paper only subscription. This is true and it would seem on a superficial level not to affect your decision-making process. But it does, the middle price option enables people to be drawn to the third. A fundamental part of the human brain is to judge things relative to their surroundings. By having the middle option alongside the first and third it will increase the sales of the third option by 260%, rather than only giving the first and third option. (Ariely, 2009) So, the middle 'decoy' may seem careless and almost a mistake, but it is secretly manipulating the decision-making process and preventing rational decisions of many individuals. This evidence demonstrates the extent to which a nudge can manipulate every decision you are faced with, without you even knowing it is happening. In a world where transparency within companies is becoming ever more important this degree of manipulation will surely be deemed mind control by many and can only bring widespread scepticism in the eyes of the public.

CONCLUSION

The significance and prevalence of nudges within our lives will only grow with time and the careful principles that govern a nudge will become more important in an increasingly complex world driven by technology. On balance, it is my opinion that a nudge is a legitimate form of marketing. The fact that nudges must only be an offer and the customer must always obtain the right to choose any option clearly indicates a principal underpinning marketing, whereas if it were to be mind control this principle would have been sacrificed. In addition, because of the second characteristic that nudges share with marketing, that of being a beneficial trade for the customer stemming from the philosophy of paternalism, it indicates that the qualities which underpin nudges more closely resemble that of marketing. Although on many levels a nudge could be deemed mind control, due to the unethical nature of its use of weaknesses within the human cognitive system and its manipulative qualities of not informing the recipient of its motives. The fact that despite influencing decisions the individual can always choose freely, outweighs the unethical nature by which nudges work. Due to the fact that companies are intrinsically unable not to nudge and the lack of sinister motives the manipulative qualities of nudges are lessened. So as long as the freedom of choice is preserved, and the nudge is beneficial then it must be seen as a legitimate marketing technique rather than subversive mind control.

What is Sentience and to what extent can this concept be applied to AI (Artificial Intelligence)?

Matt Scully, Upper Sixth

Commonly understood to mean the ability to feel, perceive, understand, and experience subjectively, sentience is a term most readily associated with humanity. (Juan Carlos Marvizon, 2019) However, since the inception of AI the question was asked as to whether sentience could also be applied to machines. Can a machine think and feel much the same as a human? I believe the answer to this is yes and that machines may even surpass our own capacity for these qualities. It is necessary therefore to break down the components of sentience and understand what it means to experience something subjectively, to perceive and understand and perhaps most importantly to feel emotion.

The argument for sentience in machines was first iterated most notably in 1950 by Alan Turing who proposed that if a machine acted as if it was sentient then it should be considered sentient. To justify this Turing proposed the imitation game (Britannica, 2020) wherein two participants, one a machine the other a human, were subjected to questions by a remote human interrogator who must within a fixed time frame decide which is a machine. From Turing's perspective a machine which wins the game and tricks the interrogator must therefore itself be sentient. However, critics of the argument would highlight that a machine that can do this is not necessarily exhibiting original thought merely repeating what it has been fed albeit sophisticatedly. This led to multiple counter arguments most notably one by John Searle.

Searle proposed the experiment of the Chinese Room (Cole 2020) in which he theorized the idea of a room with a man in it packed with containers of Chinese symbols alongside the necessary instructions to manipulate these symbols so as to be able to see which symbol to show upon seeing another. Outside of this room were people who would send in Chinese symbols to which the man would send back the corresponding symbols the instructions told him to. Unbeknownst to him he was answering questions given to him by the people in Chinese and thus he has passed the Turing Test for understanding Chinese despite not understanding a word of Chinese. Searle's argument serves as a manual analogy of a computer with its data base of symbols and instructive programs and highlights how a computer's ability to give an output that is correct does not imply it has any understanding of the subject material. Searle demonstrates here his belief that sentience is derived from the processing that takes place between input and output and that synthetic replication is not a valid candidate. A sentient being would process an input and through perception and knowledge of the subject at hand provide an output. Our sentience as humans is reliant on being conscious of the process that leads us to that output. This gives rise to the question of circumstances where humans act without thinking for example relinquishing your grip on a hot pan. The action was taken without thought and you only understand and process what happened after the fact. Logic follows then that by this definition one could not have been sentient at that moment as the process between the input and output here was not an understanding and perception of the input but instead a predetermined response much the same as with a computer.

Considering that as humans we do have a series of pre-set responses one could compare the brain to the concept of a biological computer with its major difference being that of original thought. The brain could indeed be synthesized with its collection of neurons and synapses artificially with the ability to process knowledge however it would need to create an original thought itself to distinguish itself from a regular machine. As humans our sentience is not based on definitions of living and non-living but instead independent thought and programmed thought. The idea of this originality of thought could be portrayed like this in the following example. There is a comedy writer and a machine which has been programmed to write comedy and they both write a script for a new tv show. The machine writes its script based on the entire archive of every comedy programme ever made and the writer has written it influenced by what he has seen but also by his own experiences. Whilst both shows created are comedy only the latter one will display any originality as it has been based off subjective experience. The former whilst looking original due to the vast number of sources it has taken from is just a cut and paste copy and brings nothing new to the sector of comedy. It is here that the aspect of subjective experience is demonstrated illustrating that a core part of sentience lies within the ability to react to an event based off the formation of a thought at that time rather than one that is there from the start and is called upon when needed. This would be irrelevant to Turing's definition of sentience however Searle illustrates its flaws plainly enough for it to be clear that Searle's concepts of understanding and original thought are the more accurate way to define sentience.

It is necessary to analyse the concept of understanding and its applications within the meaning of sentience to determine if this definition of sentience could apply to AI and if we would even be able to realise it. There is an upper limit of understanding to all beings, a dog cannot understand the intricacies of quantum mechanics and a bird cannot understand and interpret the entire language of Spanish. The same is therefore true of humans. We have over the millennia raised the upper limit of this understanding from understanding how to make fire to understanding how to clone DNA for example. However, currently there is so much of the universe we cannot comprehend due to this limitation. Naturally as time progresses, we will most likely continue to raise the limit however there may come a time where we no longer can. It is entirely plausible therefore that AI could reach a plane of sentience perhaps even greater than our own and we would not even realise due to our inability to comprehend it as AI's upper limit of comprehension surpasses our own. It is at this point where AI may have understood the nature of reality but is unable to explain it to us as we simply cannot grasp it. In fact, if the basis of being human is dependent on our sentience and ability to feel and perceive, arguably an AI that can demonstrate those qualities so greatly to the point it goes beyond our own capacity then is that AI more human?

Having addressed both the matters of perception, understanding and subjective experience it leaves arguably one of the hardest components to analyse, the concept of feeling. It is often argued that whilst reliant on sentience, the ability to feel emotions is the core of being human and that which divides us from every other species. So, do feelings and emotions that have been artificially replicated

still count as feeling? This brings us back to Turing's argument that much the same with sentience if a machine can display feelings real enough to convince a human it can feel then it can feel. It is also important to note that feelings are created by the brain and additional chemicals all of which theoretically could be replicated. It follows therefore that there is considerable similarity between the two brains with the only exception being one was artificially synthesised.

Theoretically then we have created an artificial intelligence which can perceive, feel, and understand. For all intents and purposes this creation is identical to a human with the exception of being artificially created rather than conceived. So, it stands that this creation must be at least as sentient as a human. Considering that an AI could be more sentient as discussed earlier it also suggests that sentience is a scale rather than a simple yes or no definition. To this same extent we should consider self-awareness as a scale as well. The human experience and thus subjective experience are defined by our knowledge of our mortality which influences the decisions we make. To this extent even if a machine possessed both self-awareness and sentience it could not have these qualities in the quantity humans do as even if it were to understand its existence it is not confined by mortality. This would suggest that whilst this machine could still be on the scale of sentience it would be impossible for it to share the same space as humans on that scale. Further to this self-awareness lends itself most easily to the scale model as to be completely self-aware you must understand absolutely everything and all its implications. It is therefore justifiable to state it is a scale as humans are self-aware despite not knowing everything and it would be illogical to state that we as humans who are aware of our limited lifespan and a sizeable portion of our environment are not self-aware. We must then fall onto a scale much the same as with sentience where we can be either more or less self-aware.

Considering a machine can emulate all the qualities that makes a human to the point where it is indistinguishable from a human the question no longer revolves around its capacity for sentience but rather if this sentience is something we are capable of comprehending. We understand our own sentience as we experience it. This specific level of sentience is exclusive to the human experience however the human experience can be emulated. The concept of mortality can be implemented into artificial intelligence to such a degree that its level of sentience is identical to a human. This would of course be a level of sentience we could understand however as by its definition we could not understand something that exceeds our sentience as it is a self-limiting factor. This gives rise to the question that we may never notice a machine that has obtained sentience as we simply lack the capacity to understand it. It is estimated by scientific revolutionaries throughout the world for this take 30-40 years in actuality and theoretically I believe it is certainly possible. The only difference distinguishable is the concept of the real and the simulated. However, the two concepts blur into one as the simulations becomes so advanced there is no difference in any way, shape, or form. If something is simulated so accurately that it is indistinguishable in any way, I would argue much the same as Turing would, that it is real. The argument of simulation and reality only applies to the point where a machine is equally or less sentient than us. After it surpasses this point it becomes real in the definition that it cannot simulate as it is the first to achieve this point. Its experience is undoubtably real at this point which is highlighted by the fact we would not understand it confirming it cannot be emulating humanity at that point.

In closing it is apparent that sentience cannot be applied as a quality but instead viewed as a scale. Machines have the capacity for sentience undoubtably and will easily demonstrate this but that this sentience can ascend higher on the scale then we could possibly understand. The core components of sentience, the ability to feel, perceive and understand as well as subjectively experience are all components a machine could incorporate. First emulating them and then surpassing our capacity for these components and climbing the scale, no longer needing to simulate these components as its upper limit of knowledge climbs far beyond what ours may ever reach. It is impossible to understand what form this sentience would take and how it would operate but it is possible to acknowledge that one day we will be surpassed as the most sentient and self-aware life forms on this planet and artificial intelligence is the most likely candidate to do it.

BIBLIOGRAPHY

- Britannica, T. E. (2020, May 07). *The Turing Test.* Retrieved June 26, 2020, from Encyclopedia Britannica: https://www.britannica.com/technology/Turing-test
- Cole, D. (2020, February 20). *The Chinese Room.* Retrieved June 26, 2020, from Stanford Encyclopedia of Philosophy: https://plato.stanford.edu/entries/chinese-room/
- Grey, C. (2014, August 13). Humans Need Not Apply. Retrieved June 27, 2020, from https://www.youtube.com/watch?v=7Pq-S557XQU
- Juan Carlos Marvizon, P. A. (2019, August 26). *What is Sentience.* Retrieved June 26, 2020, from Speaking of Research: https://speakingofresearch.com/2019/08/26/what-is-sentience/
- LEMMiNO. (2016, December 22). Artificial Intelligence. Retrieved June 27, 2020, from https://www.youtube.com/watch?v=5J5bDQHQR1g
- Murray, A. (2018, August 10). *When machines become sentient we will have to consider them an intelligent life form.* Retrieved June 26, 2020, from The London School of Economics and Political Science: https://blogs.lse.ac.uk/businessreview/2018/08/10/when-machines-become-sentient-we-will-have-to-consider-them-an-intelligent-life-form/
- PhD, M. B. (2016, July 15). *Animals as Persons: Can We Scale Intelligence or Sentience.* Retrieved June 29, 2020, from Psychology Today: https://www.psychologytoday.com/us/blog/animal-emotions/201607/animals-persons-can-we-scale-intelligence-or-sentience
- Tegmark, M. (2018). *Life 3.0.* Penguin Books.

The RGS Explore Competition Winner

Noah Campbell, Upper Sixth

YOU ARE A PRODUCER WORKING FOR THE BBC. YOU'VE BEEN ASKED TO PRODUCE MUSIC FOR AN UPCOMING DOCUMENTARY ON THE BRITISH RAJ, WHICH WILL TELL THE HISTORY OF THE BRITISH COLONISATION OF THE INDIAN SUBCONTINENT TO LARGE AUDIENCES AROUND THE COUNTRY

British rule over India is possibly the most controversial topic of the history of the British Empire. Some argue that Britain developed India's economy, legal and administrative system, and its role in world politics. However, India was as developed as Britain in the 1700s when the British arrived and if there were any benefits from British rule, these went to the British and most Indians gained little. The British Raj was the period of British rule between 1858 and 1947. During these times, 20% of British exports went to India and Britain received a huge amount of goods from India which were then exported on. The Indian army was Britain's greatest resource and around 40% of India's wealth was spent on this which was used by Britain, including in the First and Second World Wars. Britain's mercenary venture into India was fuelled by a lust for power by the expanding British Empire.

My assignment is to produce music for the following sections of a BBC documentary on the British Raj: The First War of Independence, Durbar of 1911, Gandhi's return to India in 1915, The Jallianwala Bagh massacre 1919 and Independence 1947.

The music used to depict non-Western cultures in BBC documentaries is often mildly appropriative, Orientalising and, as such, insensitive. Therefore, I will make the following measures to avoid insensitive content:

- Use of non-Oriental music where Oriental music is unnecessary and Oriental music will only be used for factual descriptions of Indian culture.
- Use of dynamics and speed of merging pieces to create an atmospheric effect but not used to present obvious emotive influence on viewers.
- Use of silence in certain aspects of the documentary as another way of depicting impact of conflict and loss of life, tradition, landscape etc.
- Use of music that does not originate from media sources such as movies, as recognisable music might diminish the importance of facts presented in the documentary.

THE INDIAN REBELLION/FIRST WAR OF INDEPENDENCE 1857

The First War of Independence occurred when Indian troops (Sepoys) serving in the army began a rebellion for several grievances: racial discrimination within the army, uniform regulations prohibited Hindu and Muslim Sepoys from wearing religious marks and required them to trim their facial hair, animal fat was introduced to grease weaponry - a taboo for Hindus and Muslims, the British influenced religion in India and the land revenue system took land from Indian natives. The war was brutal, bloody, lasted 18 months and was unsuccessful for the Indians.

The opening for the documentary on India's rough and desperate fight for independence, depicts scenes of India before the British arrival. Raman Kalyan - Ahir Bhairav [Sunrise in India] accompanies scenes representing the tranquil and beautiful essence of an unconfined India. The use of the traditional Indian bamboo flute conveys a sentiment of peace and its use in Hindustani classical music links with the Indian Hindu tradition of meditation, an intrinsic part of Indian culture and religion throughout time.

The First War of Independence scenes are accompanied by 'Smetana- Ma Vlast: Vltava' as India's first attempt at freedom from the clutches of the British Empire is explained. 'Smetana- Ma Vlast: Vltava' describes the unrest of the Indian people through its contrapuntal wind melodies and symbolises the rapid uprising of the Indian rebellion forces as their hope for change increases. I have a personal link with this piece having played it in an orchestra. Whilst performing it, I experienced feelings of inspiration and hope. Emotions I imagine perfectly define those driving a rebellion. The modulations between major and minor sequences of the counterpoint flutes and clarinets form uncertain yet beautiful imagery, depicting ideals of hope and a plentiful future.

In stark contrast to the calm and peaceful nature of Raman Kalyan - Ahir Bhairav [Sunrise in India], 'Mars' from Gustav Holst's The Planets Suite accompanies scenes of a broken and conflicted India following the unsuccessful attempt at independence. The end of Mars represents a crushing defeat as the repeated homorhythmic booming sequence ends with one harmonious yet sinister chord. India and its profitable resources remained under British rule as the British Empire sought to become the world's most powerful force. Here, silence accompanies a scene of dead Indian soldiers.

THE DURBAR OF 1911

Following the First War of Independence, the British Crown governed India and Britain benefitted from India's profitable resources. Although it is believed the Raj intended to increase Indian participation in governance, the powerlessness of Indians to do this without British consent led to discontent. Overall, the British were invasive to the Indian infrastructure, traditions, and culture. The Durbar of 1911 represented British insensitivity. In a highly controversial celebration of the rise of British Monarchy in India, it marked the coronation of King George V as King-Emperor of India. The Durbar was fashioned and timed in poor taste and amidst the pomp and splendour, it was also announced that Delhi would become the capital of British India and not Kolkata (Calcutta).

The momentous occasion when medals were distributed to every serving army member in attendance, despite the controversy of the event is accompanied by Fanfare for St. Edmundsbury by Benjamin Britten. The Fanfare reflects the spectacle of the occasion perfectly. The silence surrounding the unaccompanied

sound of the trumpets symbolises the dissonance of the Indian people and alludes not only to British victory but also to India's defeat.

O sacrum convivium by Messiaen then provides a sacred tone during the appointment of the British monarch as King-Emperor of India, whilst also offering an uneasy and conflicted tonality to represent the discontent associated with his position. The haunting melody of the piece reflects on the loss of Indian lives during the First War of Independence, which led to the British Raj.

The music switches to Jupiter from the Planets Suites after the coronation to emphasise the thoughtlessness of the occasion. This music stresses the ease with which India was taken, and the power Britain exerted over the country; strong enough to throw such a disrespectful and tastelessly joyful ceremony. The jovial and powerful character of the music reflects the selfish and hungry acts of the British Empire.

Next, Miserere Mei, Deus by Allegri is introduced in this section and then again, later to symbolise betrayal at the Jallianwala Bagh massacre. Its macabre tone represents the commitment of sin and the need for God's forgiveness. Its use at this point of the documentary is to subtly link it with an act of wrongdoing and pre-empt that same association with betrayal later. Here, it should be unrecognisable in this purpose but on reflection, add musical content.

GANDHI'S RETURN TO INDIA IN 1915

Mahatma Gandhi, leader of India's non-violent independence movement against British rule advocated for the civil rights of Indians. Respected as a nationalist, theorist, and organiser in India, he was invited by Gopal Krishna Gokhale to join the Indian National Movement against the British Rule. Gandhi's return to India in January 1915 was a pivotal point; one when the voice of the Indian people started to be heard. He created the civil disobedience movement and defied the government by breaking with their laws to undermine their monopoly on India's valuable products. Gandhi was instrumental in the freedom of India movement. He gave Indians hope of a brighter future and one of freedom and independence. Gandhi's idolisation as a successful embodiment of the Indian ideology was a huge factor in his success. He was extremely influential and effective.

Gandhi's portrayal in the documentary is accompanied by multiple layered samples of Gandhi quotes, Gandhi media reports and historical references. These spoken samples sound over traditional Indian Dhrupad music. The collation of disordered media clips signifies Gandhi's influential power and his ability to unify the people of India. Dhrupad, the ancient form of the Hindustani music, is played throughout India and the word Dhrupad is derived from Dhruva, meaning immovable and permanent. This refers to India's strength and its eternal culture. The heavy and majestic style of Dhrupad emanates heroic character, an appropriate accompaniment to Gandhi's role in India's journey to independence. Dhrupad was considered a forgotten artform and due to a loss of royal/aristocratic patronage, lived through only a few families who practised the tradition. Therefore, the Dhrupad music used in the documentary will be recorded at an open concert platform where traditional music styles are performed.

THE JALLIANWALA BAGH MASSACRE 1919

On 13th April 1919, British and Gurkha troops massacred at least 379 unarmed Indians at a meeting in Jallianwala Bagh, Amritsar. Most killed were Indian nationalists meeting to protest the forced conscription of Indian soldiers and the heavy war tax imposed by the British. Brigadier General Dyer banned all gatherings in the city. However, on the day of the Sikh festival, tens of thousands of people gathered from surrounding villages to attend. Unaware of the ban, many met at Jallianwala Bagh where the nationalist demonstrators were also meeting. Dyer's troops surrounded the park and opened fire on the crowd. Several hundred were killed and more than a thousand injured. Fire continued until all the ammunition was exhausted. During WWI, Gandhi supported the British, hoping to win partial independence for India, but after this massacre, he accepted nothing less than full independence. Little remorse has been shown by the British army since and the subject is still a fragile matter to many people who lost family and friends on that day.

To properly remember the loss caused by The Jallianwala Bagh massacre and its devastating effect on India and its people, Miserere Mei, Deus by Allegri, is repeated and accompanies the documentary again but in the context of the massacre. The minor tonality mourns the tragic event, and the lyrics reiterate to viewers the commitment of a sin and the need to ask for forgiveness. Its macabre character reflects the irreversible damage inflicted on India and the chorus of voices in the selected recording (made within a large, enclosed space) creates a ghostly echo to the music. This, within the context of the massacre builds on its previous use to convey a message of the confusion and terror of those caught in the crossfire.

In the aftermath of the massacre, the fourth movement from No. 6 by Pyotr Ilyich Tchaikovsky plays. This reflects loneliness, regret, and resignation of those lost. The piece portrays Tchaikovsky's own melancholy and longing which also mirrors India's yearning for the peaceful and independent country it once was. The emotional variation in dynamics allows for the feelings of anger and despair to appear in the music, channelling the frustration Indians had with the irrational actions of the British Army.

INDEPENDENCE 1947

Before British intervention, India was flourishing but just prior to British withdrawal in 1947, India was in a terrible shape. It was deeply divided along religious lines and as independence grew closer, violence began between Muslims and Hindus. The British withdrew in 1947 and India was partitioned into two independent countries - India (mostly Hindu) and Pakistan (mostly Muslim). Around two million people were displaced, fleeing to areas where they would not be a minority. India's fight for independence was long and driven by their need for freedom and choice. They wanted their own government and to be free of British rule.

To simulate the bittersweet feeling of relief, 'Clair de Lune' by Claude Debussy, accompanies this section of the documentary. This piece perfectly projects feelings of calm and relief, but with an air of mourning. Joy and hope echo within the melody, providing a quality of reminiscence. This accompanies the documentary's remembrance of the struggles endured and the lives lost during India's fight for independence.

Finally, Raman Kalyan - Ahir Bhairav [Sunrise in India], used at the start of the documentary, is repeated; a reminder that India's culture is still alive. Here it plays as images of the changed state of India are presented, including poverty, beauty of Indian culture, and the new political system with their appointed Prime Minister.

RGS EXPLORE

CREATE YOUR OWN LANGUAGE, DEVISE A STRATEGY FOR GREENER HEATING, STUDY THE SOCIAL LIVES OF YOUR FAMILY AND FRIENDS!

RGS GUILDFORD

Anything is possible when you take the time to explore... Engage with subjects beyond the school curriculum

www.rgs-guildford.co.uk

@RGSGuildford

RGS Guildford

Registered Charity No. 312028