Assignment 2 - Project Proposal [2198 words excluding slide titles and references]

[Slide 1: Project Title]

This is my presentation for assignment 2 of Research Methods and Professional Practice, and it looks at the implementation of Machine Learning Tools and Techniques in Medication Safety Analysis, where I think it can do a lot of good in identifying and predicting problems that can be experienced by patients undergoing medication.

[Slide 2: Outline of Presentation]

First, I will talk about the contribution of this project to the field of machine learning in the detection of medication side effects, called adverse drug reactions, or ADRs. Different types of medical records will be discussed, the problems with paper-based systems and the difficulty of ingesting data from multiple sources. Then I will present the research questions that the project will be designed to answer. There are three of them, namely "How can machine learning detect and predict adverse drug reactions?", "What are the most suitable types of machine learning model?" and "Can social media data be combined with electronic health records?". There will be a slide that builds on the research questions, about the aims of the project, which are identifying the best ways to detect and predict ADRs, and the specific objectives like describing the development, evaluation and application of machine models. We'll take a look at the papers that were reviewed so far in preparation for the project, and give them a critical appraisal, before going on to talk about the methodology, development strategy and research design of this project, which will be based on established principles. Areas covered will include choice of data source types, choice of machine learning types, the training and evaluation of models, and benchmarking against mainstream systems already in use. Since this project describes medical research, there will be an ethical dimension to it, and I shall discuss this as it applies to patient privacy, data security, balance (by which I mean lack of bias) and transparency. The project will generate several artefacts in the form of examples of input data, anonymised of course, examples of source code for the models that could be used, visualisations and a final report. After

showing a timeline of the proposed project activities, I shall present the project's conclusions, or at least the anticipated headings, as it's a bit premature to conclude anything before the project has even started.

[Slide 3: Significance/Contribution to the Field of Medication]

The use of Machine Learning in medicine is a rapidly expanding field of research, and the contribution I hope this project will make is going to be an appraisal of this progress so far, along with suggestions for future research projects. There have been several surveys in the recent past, some of which will appear on the literature review slide in a couple of minutes. These have been limited in scope and have failed to integrate different sources of data into a meta-analysis bracket Wang 2023 close bracket.

The field under study is that of ADR, adverse drug reactions, and there is a serious problem with underreporting of these harmful side-effects of medicines, with only about six percent of ADRs being reported bracket Hazell & Shakir 2005 close bracket. This is happening because of the low detection rate, which, in turn is because many patient records are handled by manual systems, with only 45 percent of National Health Service records being fully electronic bracket Best 2023 close bracket, and the problems with manual reporting are the inconsistency of the data and the delays in getting them back to a central node where they can be aggregated.

Despite the problems highlighted by the three papers I just cited, Machine Learning promises to help detect and predict ADRs, especially if it can be fed data from electronic health records, or EHRs, as well as from social media bracket Dey et al., 2018 close bracket.

[Slide 4: Research Questions]

Three main research questions will be posed and answered by this project. Firstly, can machine learning detect and predict adverse drug reactions, and if so, how? Secondly, can social media data be combined with EHRs to produce comprehensive datasets which produce self-consistent results? Lastly, what are the most suitable types of machine learning models for this kind of task?

[Slide 5: Aims and Objectives]

Those research questions lead us to the aims and objectives of the project. There are two main aims here, to be able to detect adverse drug reactions using machine learning, and then to be able to predict incidences of adverse drug reactions using machine learning. The two are very different aims and it may emerge that they will require very different machine learning models. In order to achieve the aims, a number of objectives will be set, currently four but this is likely to grow as unforeseen tasks need to be completed. The first of the four identified so far is the assessment of the feasibility of converting data mined from social media into a format that can be ingested by a machine learning model. This will involve natural language processing, or NLP, to reduce this information to basic facts, then some feature engineering to prepare it for ingestion bracket Anjali & Ravi Kumar, 2022 close bracket. The second objective is to select and describe appropriate machine learning models and to describe how to train them with the converted social media data combined with electronic health records, which will need less feature engineering bracket Nguyen et al., 2021 close bracket. Then we will have to describe testing each model with data held back for the purpose, and how to evaluate them against metrics to be agreed on bracket Dey et al. 2018 close bracket. The final objective will be to recommend how to apply the winning model to real clinical settings.

[Slide 6: Key Literature Related to the Project]

From many papers about using machine learning to detect and predict adverse drug reactions, I selected the six shown here as the basis of my understanding of this field. I have shortened their titles to fit them on this slide but their full details are on the references slides. On the timeline slide you will see I have allocated a whole month for further literature review. For now, the project will draw on these six papers, the first three of which have an emphasis on ADR detection, while the last three concentrate on prediction, despite the title of the one by Hu et al., bracket 2024 close bracket, which mentions prediction but not detection - relying on titles is not enough to characterise papers. The first of the papers to concentrate on detection is Anjali & Ravi Kumar bracket 2022 close bracket, which sets

the scene by talking about the use of natural language processing to extract data from social media and put it into a form compatible with electronic health records so it can be combined with EHR data for ingestion into machine learning models that will identify ADRs. Models described include Support Vector Machines (SVMs) for text classification to identify ADR-related posts, random forest (RF) classification into ADR and non-ADR records. The second is Hu et al bracket 2024 close bracket. This reviews several other papers and makes clear that there is a need for the standardisation of data formats in electronic health records to enable ingestion to take place at scale. Even with EHRs there are inconsistencies in the way illnesses are coded, with different standards such as ICD-10, MedDRA, SNOMED CT, and WHO-ART, and they also include some data which is in natural language without any encoding. McMaster et al. bracket 2019 close bracket, the third of the papers on ADR detection, addresses this problem by advocating using ICD-10 codes in patient records, as these are already heading for universality and are extensible for enhanced granularity in describing conditions in more detail. In my own opinion, if the four standards just mentioned have roughly one to one equivalences, there should not be a problem combining medical records with different codings, but this may not prove to be feasible. The first of the papers that major on the power of machine learning to predict adverse drug reactions is Nguyen et al., bracket 2021 close bracket. This one discusses the use of support vector machines (SVMs) and k-nearest neighbours (KNN) for classifying drugs according to their risk of causing the ADRs identified using the detection models, and it goes further than this, to consider the possible analysis of ADR mechanisms. The second paper about prediction, Islam et al. bracket 2018 close bracket, discusses the use of Federal Drug Administration data to develop machine learning to predict not only the occurrence of ADRs but also their severity, which, if it can be done, would be invaluable in identifying those most at risk. Support Vector Machines for classifying low and high risks, Random Forests for classifying mild, moderate, severe and fatal risks, K-Nearest Neighbors for case comparisons, Logistic Regression for predicting severity on continuous scales, and Decision Trees for classifying severity into a hierarchy. The last paper, Dey et al. bracket 2018 close bracket, is the most optimistic about ADR prediction, perhaps because it goes into deep learning techniques, which it claims will be highly accurate, efficient, scalable and adaptable. The paper advocates using Multi-Layer Perceptron (MLP) for ADR classification, Recurrent Neural

Networks (RNNs) for analysing time-series data, and Convolutional Neural Networks (CNNs) for automatically interpreting biomedical images.

[Slide 7: Methodology/Development Strategy/Research Design]

The methodology used in this study will be that of a systematic literature review (SLR) to widen and update previous such reviews so as to present the state of the art. It will follow methodologies suitable for computing and information system projects, as described by Dawson open bracket 2015 close bracket, chapter 3. Data sources will be characterised, including the advantages and disadvantages of structured data like EHRs, ADR databases, and unstructured data such as social media, doctors' notes etc. The problems of standardisation and potential bias will be examined bracket Hu et al., 2024 close bracket. The ways of preprocessing raw data into machine learning ingestible formats will be explored, with emphasis on social media data conversion using natural language processing techniques like word embeddings, sentiment analysis and named entity recognition (NER) bracket Anjali & Ravi Kumar, 2022 close bracket. Much emphasis will be given to the choice of machine learning models, for example supervised learning, like SVMs and logistic regression, and also the different types of deep learning neural networks. The trade-offs between accuracy, interpretability, and computational cost will be made clear bracket Dey et al., 2018, Islam et al., 2018; McMaster et al. 2019 close bracket. I will describe splitting the data into training, validation and test sets, including cross-validation, and discuss the parameters that need to be specified in models bracket Nguyen et al., 2021 close bracket. The metrics for evaluating machine learning models will be compared, including accuracy, precision, recall and F1-score bracket Bruce et al., 2020: 219-223 close bracket. Benchmarking will consist of comparisons with existing methods, and sources will be found, probably including the Federal Drug Administration and the World Health Organisation. The trade-off between accuracy and explainability was covered in McMaster et al. bracket 2019 close bracket, and I will expand on this.

[Slide 8: Ethical Considerations and Risk Assessment]

None of these operations should be performed on real patient records without regard to the applicable ethics in these areas. Patient privacy should be respected at all levels of the data processing. Researchers should abide by the General Data Protection Regulation, or GDPR bracket European Parliament, 2016 close bracket, the Association of Computing Machinery Code of Ethics bracket 2018, Section 1.6: Privacy close bracket, and the BCS Code of Conduct bracket 2022, Section 1d close bracket. I think the safest way to ensure privacy is to use anonymised sources. Data security is closely related, and the care that must be taken to guard data is covered by the ACM Code of Ethics bracket 2018, Section 1.2: Avoid Harm close bracket, which is about preventing unauthorized access to data. It's vital to avoiding biases in datasets that could lead to misdiagnosis or unequal treatment, and this is mandated by Section 19 of the UK Equality Act bracket UK Government, 2010 close bracket, which prohibits discrimination. Finally, transparency must be maintained throughout, and this is in order to gain the trust of the healthcare sector. According to the BCS Code of Conduct bracket 2022, Section 2c close bracket, there is a duty to "explain computing solutions in a way that users can understand."

[Slide 9: Artefacts That Will Be Created]

The report could include samples of EHRs, in the form of screenshots or data in comma separated variable format. Also, screenshots of raw social media data could be included. Listings of Python code, showing how machine learning models are created, can be included. Many possible visualisations could be provided, including charts illustrating the distribution of patient record types, the workflows to create, train, test and evaluate models, decision trees, model performance etc. The final artefact will be the actual report, containing all of the above.

[Slide 10: Timeline of Proposed Activities]

The project will take seven months, one for the literature review, two each for describing data collection and model development methods, one for describing evaluation and another for the writeup.

[Slide 11: Conclusion]

I will make conclusions about the feasibility of combining data sources, ranging from dubious bracket McMaster et al., 2019 close bracket to enthusiastic bracket Dey et al., 2018 close bracket, effective machine learning models for detecting and predicting ADRs, and their applicability to clinical settings.

[Slide 12: References]

Association of Computing Machinery (2018). ACM Code of Ethics and Professional Conduct. Available at: https://www.acm.org/code-of-ethics.

Anjali, B.V. & Ravi Kumar, G.K. (2022) 'A Broad Review on Adverse Drug Reaction Detection using Social Media Data', *6th International Conference on Intelligent Computing and Control Systems (ICICCS)*. Madurai, India, 25-27 May 2022. Piscataway, NJ, USA: IEEE. DOI: doi.org/10.1109/ICICCS53718.2022.9788381

BCS, The Chartered Institute for IT (2022). Code of Conduct for BCS Members. Available at: https://www.bcs.org/media/2211/bcs-code-of-conduct.pdf

Best, J. (2023). NHS still reliant on paper patient notes and drug charts despite electronic upgrades, the BMJ finds. *BMJ*. DOI: doi.org/10.1136/bmj.p2050

Bruce P, Bruce, A. & Gedeck, P. (2019) *Practical Statistics for Data Scientists*. 2nd ed. Sebastopol, CA: O'Reilly Media Inc.

Dey, S., et al. (2018) Predicting adverse drug reactions through interpretable deep learning framework. BMC Bioinformatics (19)476. DOI: doi.org/10.1186/s12859-018-2544-0

European Parliament and Council of the European Union, 2016. Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data (General Data Protection Regulation, GDPR). *Official Journal of the European Union* L 119: 1-88. Available from: https://eur-lex.europa.eu/eli/reg/2016/679/oj [Accessed 31 Mar. 2025].

Hazell, L., & Shakir, S.A.W. (2006). Under-reporting of adverse drug reactions: A systematic review. *Drug Safety* 29(5): 385–396. DOI: doi.org/10.2165/00002018-200629050-00003

Hu, Q. et al. (2024) Machine learning to predict adverse drug events based on electronic health records: a systematic review and meta-analysis. *Journal of International Medical Research* 52(12): 1-19. DOI: dx.doi.org/10.1177/03000605241302304

Islam, T., Hussain, N., Islam, S. & Chakrabarty, A. (2018) Detecting adverse drug reaction with data mining and predicting its severity with machine learning. *2018 IEEE region 10 humanitarian technology conference.* Sri Lanka Institute of Information Technology. DOI: doi.org/10.1109/R10-HTC.2018.8629806

UK Government (2010). Equality Act 2010. Available from: https://www.legislation.gov.uk/ukpga/2010/15/contents [Accessed 2 April 2025]

Wang, R. (2023) Combining Data from Multiple Studies. *NEJM Evidence 2023* 2 (5). DOI: doi.org/10.1056/EVIDe2300066