

Sean Farrell

sean.farrell2@durham.ac.uk | seanfarr788.github.io | linkedin.com/seanoliverfarrell

Career Summary

A highly driven PhD candidate specialising in natural language processing (NLP) and big data analytics, grounded in a robust foundation of biomedical sciences. My academic and professional journey is marked by a proven record of impactful, peer-reviewed publications, successful interdisciplinary collaborations across international teams, and hands-on experience teaching NLP to diverse audiences. Driven by a passion for addressing critical public health challenges, my work combines computational innovation, technical expertise, and strong communication skills to transform data into actionable insights with meaningful public health impacts.

Skills

Natural Language Processing (NLP) | Python | R | SQL | Statistical Modelling | Deep Learning | Multi-modality | Explainability | Project Leadership | Teaching | Award winning public speaking | Big Datasets | Data Visualization | GIT project management | Technical Writing | Epidemiology | Workshop Design and Delivery

Education

PhD Natural Language Processing | September 2021 – September 2025 | Durham University

Biotechnology and Biological Sciences Research Council (BBSRC) funded to leverage NLP and deep learning to analyse over 10 million free text veterinary electronic health records from practices across the UK to understand critical public health matters around antimicrobial use, premature mortality, and disease outbreak detection. Additionally, I paved the way to more open science principles through the first introduction of veterinary EHR benchmarks and a new international standard for the sharing of clinical notes

Thesis: Natural Language Processing for Early Detection and Mitigation of Critical Public Health Threats

2:1 BSc(hons) Biomedical Sciences | September 2018 – June 2021 | University of Kent

Proficient in applying interdisciplinary knowledge across biology, medicine, and healthcare to analyse and address health-related challenges. Experience in laboratory methods in Genetics, Microbiology, Biochemistry, and Molecular Biology

Thesis: Antimicrobial usage in hospitalized SARS-CoV-2 patients and the impact on the Gut Microbiome

Work Experience

Data Science Intern | June 2024 – September 2024 | Evergreen Life

I developed a LLM pipeline to provide personalized healthcare advice, ensuring the accuracy of the information by aligning it with the Evergreen Life article repository. I also created an algorithm-driven content recommendation system to tailor advice to each individual user. My work is currently being integrated into an app used by over 1 million NHS patients.

Natural Language Processing Demonstrator | January 2021 – Present | Durham University

Teaching NLP from foundational machine learning to transformers. Facilitated diverse learning experiences tailored to both technically adept audiences (MSc Computer Science students) and less technically oriented cohorts (MBA Business Analytics students). Assisted in material and course design and coursework marking.

Undergraduate Researcher | July 2020 – October 2021 | University of Liverpool

Led research on flea prevalence patterns in companion animals, analysing electronic health records from 34,000+ pets across UK. Research findings on seasonal trends and risk factors published in the Journal of Medical and Veterinary Entomology, advancing evidence-based approaches to parasite prevention.

Undergraduate Researcher | June 2019 – January 2020 | University of Liverpool

Conducted epidemiological research into myxomatosis. First exposure to large-scale data analysis and statistical methodologies utilising R. My research was published in the Journal of Preventive Veterinary Medicine.

Customer Experience Supervisor | September 2016 – September 2021 | Sainsbury's

Publications

First Authorships

- **Farrell, S.**, Anderson, K., Noble, P.-J.M. and Al Moubayed, N. (2024). **Premature mortality analysis of 52,000 deceased cats and dogs exposes socioeconomic disparities.** *Scientific Reports*, 14(1). doi: [10.1038/s41598-024-77385-8](https://doi.org/10.1038/s41598-024-77385-8).
- **Farrell, S.***, Burton, J.*, Noble, P.-J. and Al Moubayed, N. (2024). **Explainable text-tabular models for predicting mortality risk in companion animals.** *Scientific reports*, 14(1) doi: [10.1038/s41598-024-64551-1](https://doi.org/10.1038/s41598-024-64551-1).
- **Farrell, S.**, Appleton, C., Noble, P.-J.M. and Al Moubayed, N. (2023). **PetBERT: automated ICD-11 syndromic disease coding for outbreak detection in first opinion veterinary electronic health records.** *Scientific Reports*, 13(1). doi: [10.1038/s41598-023-45155-7](https://doi.org/10.1038/s41598-023-45155-7).
- **Farrell, S.**, Bagcigil, A.F., Chaintoutis, S.C., Firth, C., Aydin, F.G., Hare, C., Maaland, M., Mateus, A., Vale, A.P., Windahl, U., Damborg, P., Timofte, D., Singleton, D.A. and Allerton, F. (2023). **A multinational survey of companion animal veterinary clinicians: How can antimicrobial stewardship guidelines be optimised for the target stakeholder?** *The Veterinary Journal*, 303, pp.106045–106045. doi: [10.1016/j.tvjl.2023.106045](https://doi.org/10.1016/j.tvjl.2023.106045).
- **Farrell, S.**, McGarry, J., Noble, P.-J.M., Pinchbeck, G.J., Cantwell, S., Radford, A.D. and Singleton, D.A. (2023). **Seasonality and other risk factors for fleas infestations in domestic dogs and cats.** *Medical and Veterinary Entomology*, 37(2). doi: [10.1111/mve.12636](https://doi.org/10.1111/mve.12636).
- **Farrell, S.**, Noble, P.-J.M., Pinchbeck, G.L., Brant, B., Caravaggi, A., Singleton, D.A. and Radford, A.D. (2020). **Seasonality and risk factors for myxomatosis in pet rabbits in Great Britain.** *Preventive Veterinary Medicine*, 176, p.104924. doi: [10.1016/j.prevetmed.2020.104924](https://doi.org/10.1016/j.prevetmed.2020.104924).

Co-authorships

- Davies, H., Nenadic, G., Alfattni, G., Arguello Casteleiro, M., Al Moubayed, N., **Farrell, S.**, Radford, A.D. and Noble, P.-J.M. (2024). **Text mining for disease surveillance in veterinary clinical data: part two, training computers to identify features in clinical text.** *Frontiers in Veterinary Science*, 11. doi: [10.3389/fvets.2024.1352726](https://doi.org/10.3389/fvets.2024.1352726).
- Davies, H., Nenadic, G., Alfattni, G., Arguello Casteleiro, M., Al Moubayed, N., **Farrell, S.**, Radford, A.D. and Noble, P.-J.M. (2024). **Text mining for disease surveillance in veterinary clinical data: part one, the language of veterinary clinical records and searching for words.** *Frontiers in veterinary science*, 11. doi: [10.3389/fvets.2024.1352239](https://doi.org/10.3389/fvets.2024.1352239).
- Fins, I.S., Davies, H., **Farrell, S.**, Torres, J.R., Pinchbeck, G., Radford, A.D. and Noble, P.-J.M. (2023). **Evaluating ChatGPT text mining of clinical records for companion animal obesity monitoring.** *The Veterinary Record*, 194(3). doi: [10.1002/vetr.3669](https://doi.org/10.1002/vetr.3669).
- Smith, S.L., Anderson, E.R., Cansado-Utrilla, C., Prince, T., **Farrell, S.**, Brant, B., Smyth, S., Noble, P.-J.M., Pinchbeck, G.L., Marshall, N., Roberts, L., Hughes, G.L., Radford, A.D. and Patterson, E.I. (2021). **SARS-CoV-2 neutralising antibodies in dogs and cats in the United Kingdom.** *Current Research in Virological Science*, 2, p.100011. doi: [10.1016/j.crviro.2021.100011](https://doi.org/10.1016/j.crviro.2021.100011).

In Review

- **Farrell, S.**, Radford, A.D., Noble, P.-J.M. & Al Moubayed, N. **PetEVAL: A veterinary free text electronic health records benchmark for the anonymisation and identification of disease patterns.** [In review]
- **Farrell, S.**, Singleton, D.A. Radford, A.D., Pinchbeck G., Noble, P.-J.M. & Al Moubayed, N. **Automated Disease Classification of Veterinary Clinical Narratives for Antimicrobial Stewardship Guideline Monitoring.** [In review]
- Noble, P.-J.M., **Farrell, S.**, Al Moubayed, N. & Radford, A.D., **Comprehensive representation of health-related phenotypes in one million dogs using topic modelling of electronic health records.** [In review]

*Equal Contribution