

Identifying Military Veterans in a Clinical Research Database using Natural Language Processing

Leightley, D^{1*}, Mark, MK¹, Pernet, D¹, Murphy, D², Fear, TN¹, and Stevelink, SAM¹

¹King's Centre for Military Health Research

²Combat Stress

Background

There is a lack of quantitative evidence concerning United Kingdom veterans who access secondary mental health care. This is mainly due to a person's veteran status not being routinely collected when they enter the health care system.

Main Aim

The study aimed to develop an NLP approach for identifying veterans accessing secondary mental health care services using National Health Service electronic health records.

Methods

Veterans were identified using the South London and Maudsley Biomedical Research Centre (SLaM) case register – a database holding secondary mental health care electronic records for the South London and Maudsley National Health Service Trust of 300,000 patients. We developed two methods. An NLP and machine learning tool were developed to automatically evaluate personal history statements written by clinicians.

Results

This study showed that it was possible to identify veterans using the NLP and machine learning approach on a sub-set of 4,200 patients. The automatic machine learning method was able to identify 270 veterans representing an accuracy of 97.2%. It is estimated to take between 6 to 16 minutes to manually search patient history statements whereas the automatic machine learning method took only one minute to run.

Conclusion

We have shown that it is possible to identify veterans using NLP combined with machine learning. This work contributes towards the development of a more comprehensive picture of

veterans who are accessing secondary mental health care services in the UK. It represents a first step in identifying veterans from one dataset and we hope that future research can inform the possibility of deploying the methods nationally. Despite our success in the current work, the tools are tailored to the SLaM dataset and future work is needed to develop a more agnostic framework.

Funding

Forces in Mind Trust

*Corresponding Author:

Email Address: daniel.leightley@kcl.ac.uk (D Leightley)

