RESEARCH USING SAVSNET DATA

SAVSNET
TOPIC MODELING

USING AI-BASED TOPIC MODELLING TO UNDERSTAND DISEASE IN 1 MILLION DOGS

Full paper published in Journal of Big Data

1

THE PROBLEM

Understanding patterns of disease in populations is important for their health and welfare.
Historically, many studies have been small in scale.

In the electronic era we now have huge quantities of electronic health data.
But we just don't have the time to look through it for important patterns and signals of disease.

2

AIM OF STUDY

Artificial intelligence offers new ways to explore large volumes of data.

Here, we implemented an unsupervised machine learning methodology to create a representation of 1million clinical narratives collected by SAVSNET, the Small Animal Veterinary Surveillance Network.

3 METHODS

EXAMPLE NARRATIVE
owner says depressed and lethargic 2days. now vomit++. otherwise NAD. Advise re diet and revisit 1wk if nec.
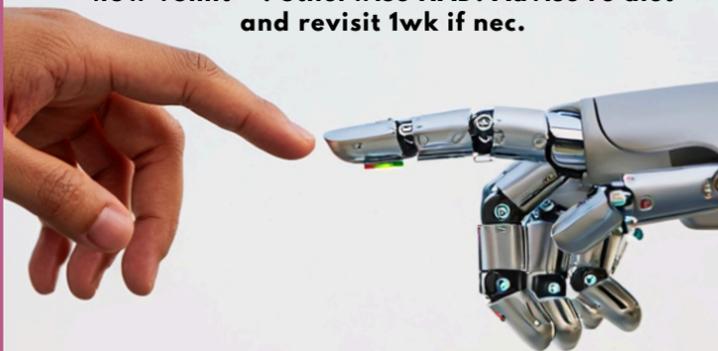
SAVSNET collects data from over 500 veterinary premises across the UK. The current study used a random sample of one million clinical records, each from a unique dog and comprising deidentified clinical narrative, the breed, age, sex and the date of consultation.
We used an AI method called **BERTTopic** to try and cluster these narratives into clinically relevant groups (or TOPICS)
These topics are discovered automatically rather than being pre-defined by a clinician

4

RESULTS - INCREASED RISK

TOPIC modelling was able to surface known diseases and identify animals at increased risk e.g.
Diabetes was more common in samoyeds, huskies and West Highland Wihite Terriers.
Thyroid disease more common in poodles & dobermanns.
Heart disease in Cavalier King Charles Spaniels.
It also identified some very seasonal patterns associated with tick bites, firework phobias and grass seeds.
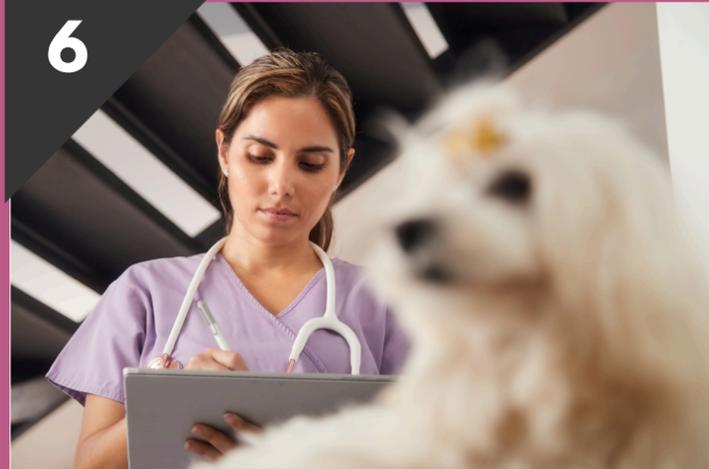
5

RESULTS - OUTBREAK DETECTION

Two topics clearly changed over time.
One was of respiratory disease in 2021.
The second was gastrointestinal disease in 2020.
This GI topic closely followed a known outbreak that occurred in the UK at that time and suggests this method could efficiently identify national outbreaks of disease.

6

THE FUTURE

Analysing large volumes of data in this way offers new opportunities to....

- make better use of the large volumes of pet health data now available.
- implement targeted disease screening.
- spot outbreaks of diseases early.

"We are hugely grateful to the veterinary practitioners and owners who share their anonymised pet health records with SAVSNET. Using these AI methods allows for a step change in our ability to use this volume of data to improve pet health."

PROF PJ NOBLE - LEAD INVESTIGATOR.

YOU CAN EXPLORE SUMMARY DATA FOR THESE 1MILLION TOPICS HERE

To see the rest of our AI based research click here