

OBJECTIVES

Explore the use of Artificial Intelligence-based Large Language Models (LLMs) with zero-shot capabilities to improve data analysis of social media posts related to canine pruritus, reducing reliance on labor-intensive manual data mark-up.



Figure 1: AI generated dog and vet using Midjourney

INTRODUCTION

Recent advances in AI-based Large Language Models (LLMs) have showcased their capability for zero-shot learning, where they perform tasks without prior explicit training (see example in Figure 1). This study applies zero-shot LLMs [1] to classify social media posts collected via Social Media Listening (SML) with the purpose of better understanding pet owner experiences with canine pruritus [2]. Typically, data gathered from social media includes unintended irrelevant content that complicates analysis, such that manual sorting of such vast amounts is impractical. This highlights the utility of zero-shot classifiers in effectively filtering relevant posts, reducing the need for exhaustive manual review.

WHAT IS ZERO-SHOT LEARNING?

Zero-shot learning in AI refers to the ability of a model to accurately perform tasks it has not been explicitly trained to do. In the context of AI, this means making accurate predictions or classifications based on unseen data or tasks, relying solely on general understanding and context.

WHAT ARE LARGE LANGUAGE MODELS?

Large Language Models (LLMs) are advanced AI systems trained on vast amounts of text data. They generate and understand natural language, enabling them to perform a wide range of language-based tasks without specific training on those tasks [3].

STANDARD MACHINE LEARNING APPROACH

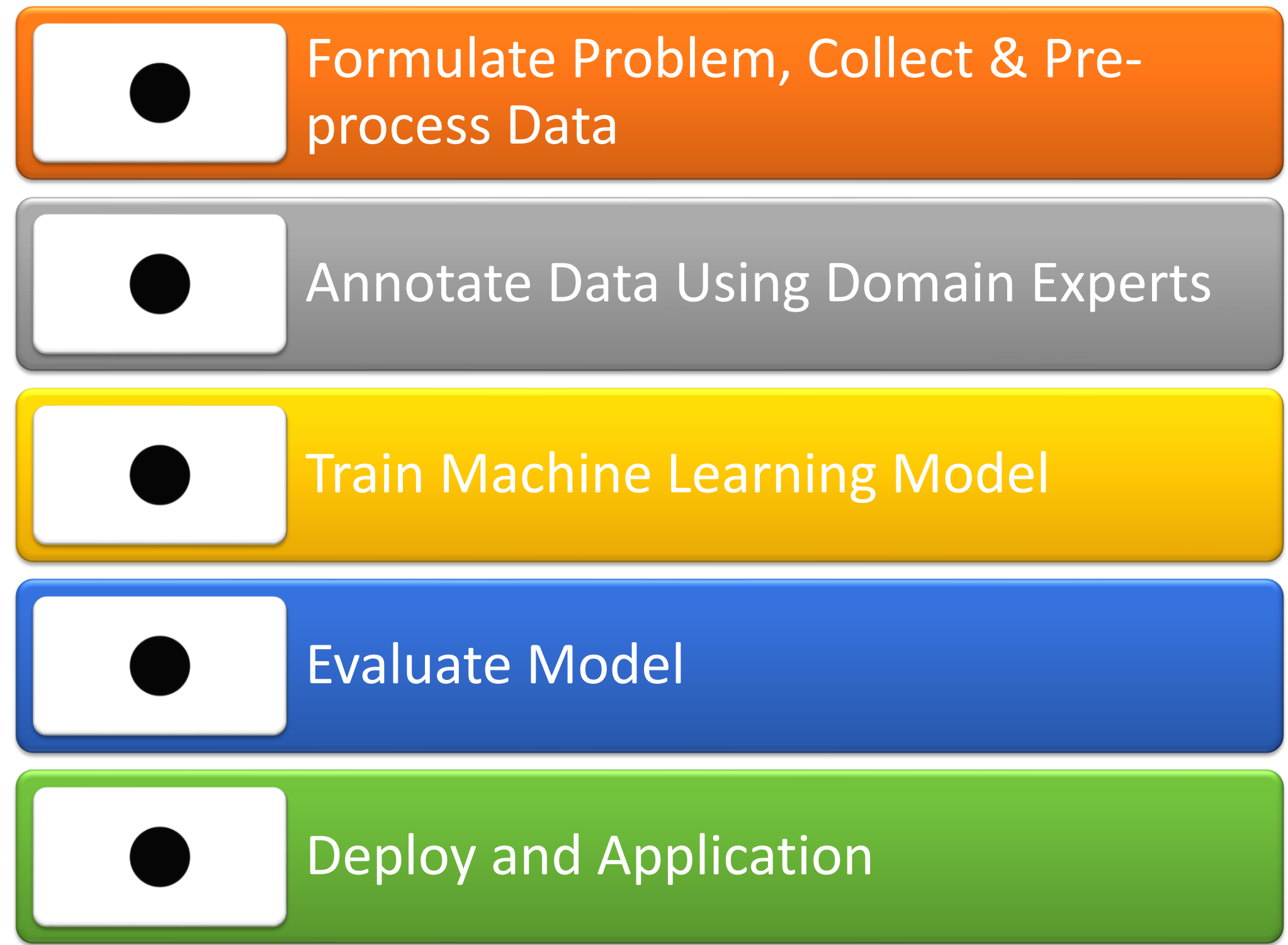


Figure 2: Flow of a typical Machine Learning pipeline

The standard supervised machine learning training process (see Figure 2) involves multiple steps, including collecting and pre-processing data, manually labelling for relevance, training a model on this dataset, and then evaluating its performance. This method, while effective, requires substantial time, expertise, and resources, making it less flexible and scalable compared to zero-shot learning approaches with LLMs. Additionally, traditional machine learning models often struggle to adapt to new or unforeseen tasks without retraining, limiting their applicability in dynamic environments where data and requirements can frequently change.

BENEFITS OF A ZERO-SHOT APPROACH

- 1. Efficiency in Deployment:** Zero-shot learning eliminates the need for collecting and labelling a new dataset for each specific task, allowing for rapid deployment and adaptation to new tasks without additional training.
- 2. Cost-Effectiveness:** By removing the need for extensive data labelling, zero-shot learning significantly reduces the costs associated with manual annotation, which is often time-consuming and labour-intensive.
- 3. Scalability:** LLMs can generalise across various tasks using zero-shot learning, making it easier to scale solutions across different domains or problems without retraining the model from scratch.
- 4. Flexibility in Application:** Zero-shot capabilities allow LLMs to handle unexpected or novel scenarios that were not covered in the training data, providing a level of versatility that traditional ML often lacks.

METHOD

SML posts from sources including but not limited to Reddit and X Corp (previously Twitter) were collected based on keywords chosen by veterinary dermatology experts for insights into canine pruritus. A domain expert manually marked-up 1560 relevant and 1477 irrelevant (total n=3037) collected posts, after a data cleaning process with duplicates removed. A descriptive text prompt defining the downstream task was provided to a GPT-3.5 LLM accompanied with each collected post. The model iteratively classified each post as either “relevant” or “irrelevant” according to the provided description. Adjusted hyperparameters of GPT-3.5 included a temperature set to zero and number of maximum output tokens set to five, to limit the output generated response and ease post-processing.

RESULTS

Our method yielded results of 0.73, 0.87 and 0.80 for sensitivity, precision and F1-score, respectively (1831 and 1206 posts were classified as relevant or irrelevant).

| Metric      | Value |
|-------------|-------|
| Sensitivity | 0.73  |
| Precision   | 0.87  |
| F1-score    | 0.80  |

Table 1: Results from the zero-shot classification using an LLM

CONCLUSIONS

LLMs, even though not specifically trained for determining relevance of social media posts, exhibit effective zero-shot capabilities when classifying relevance of SML posts for this use case on the topic of canine pruritus, thus dramatically reducing the need for manual mark-up for all posts. Combining this approach with human supervision would the aid in reducing workloads whilst potentially optimising accuracy.

REACH OUT

If you are interested in our research, or have any questions related to AI in animal health, feel free to reach out:



REFERENCES

- [1] Kojima, Takeshi, et al. "Large language models are zero-shot reasoners." Advances in neural information processing systems 35 (2022): 22199-22213.
- [2] Convertino, Irma, et al. "The usefulness of listening social media for pharmacovigilance purposes: a systematic review." Expert opinion on drug safety 17.11 (2018): 1081-1093
- [3] Wei, Jason, et al. "Emergent abilities of large language models." arXiv preprint arXiv:2206.07682 (2022).

ACKNOWLEDGEMENTS

This study was wholly funded by Zoetis.