INTRODUCTION

* Topic Background

Text summarisation is the process of condensing a passage of text into a shorter version whilst retaining the necessary information in the text. This is a valuable research area since summarisation massively reduces the comprehension time of large pieces of text. Moreover, it has applications in many different domains, both public and professional: academics are required to read extensive research papers, individuals read long articles to keep up to date with the world news, and individuals read books to learn about various topics from history to science.

\medbreak

There are two approaches to text summarisation: extractive and abstractive summarisation. Extractive summarisation is a summarisation technique which focuses on selecting particular words and sentences to convey the meaning of the original text. Whereas abstractive summarisation techniques look to understand the semantics of the text before generating text to summarise what the model has learnt.

\medbreak

Extractive summarisation models are more common as most practical text summarisation models are extractive. The basic structure of these models is made up of three stages \cite{Nenkova11}: first, capturing the key aspects of the text; secondly, using these aspects to score sentences; thirdly, to create a summary using the highest scoring sentences. Abstractive summarisation is, naturally, harder, and more computationally costly to perform. The difficulty is centred on learning the semantics of the text; different texts can have many different structures and models find it difficult to learn such variety.

\medbreak

This project focuses on developing an approach that ranks proposed text summaries. It is necessary for our approach to achieve the following requirements:

\begin{itemize}

\item {It has the ability to tailor summaries to the user’s preference since there is a range of requirements that different users have.}

\item{The highest-ranked summary suitably conveys the information in the original text.}

\item {The framework should be used in an interactive setting, so rankings must be generated without extensive computational cost.

\end{itemize}

* Approaches
  + Current

From current literature, models have been proposed to capture the preference of one summary to another such as the Bradley-Terry model \cite{Bradley52} and the Thurstone -Mosteller model \cite{Thurstone27, Mosteller51}. These models provide good solutions; however, they fail to differentiate between aleatoric and epistemic uncertainty. This limits the models’ ability to determine where there is weakness in the model and lead to a reduced performance. Alternative approaches use deep learning techniques to rank passages which beat state-of-the-art performance \cite{Xu19}. However, such models are limited by their requirement of large training data which, in the context of text summarisation, comes at high cost as human annotators are required to manually produce and evaluate summaries. Moreover, these models are unable to account for user preferences which really limits the model’s ability to tailor summaries to the user.

* + Our
    - Research Aim

Our aim is to develop and evaluate a Bayesian deep learning framework for passage ranking text summaries which incorporates an active learning component to allow for user influence on the summary rankings. Typical deep learning approaches demand vast amounts of training data; however, the active learning component will minimise the number of user interactions required, whilst maintaining high-performance. During iterations within the active learning component, we use a stream-based strategy to minimise the amount of overhead processing as, for this strategy, summary instances are evaluated sequentially by the active learner as opposed to requiring a pool of unlabelled instances. As initiated by Simpson et al. \cite{Simpson19}, we will use a Bayesian optimisation acquisition function to determine if an unlabelled instance should be queried by the oracle. We also aim to use Monte Carlo Dropout \cite{Gal15} to approximate the posterior distribution across the model weights and calibrate our model.

\medbreak

Within our framework development and evaluation, we wish to establish if the proposed Bayesian deep learning approach provide a sufficient passage ranking solution in comparison to a classical deep learning model. Moreover, we aim to determine if a stream-based active learning strategy is appropriate for such a problem.

* + - Challenges

The central challenge within the project is effectively combining a Bayesian deep learning model with an active learning component. Firstly, it is a concern as to whether stream-based learning is an appropriate active learning strategy for this task since there is minimal current documentation; pool-based active learning the most common strategy used \cite{Settles09}. Secondly, consideration needs to be made with regards to the number of user interactions. It is necessary for the model to be interactive; thus, it is important to explore the number of interactions that are required and if this is a reasonable level for an interactive setting.

* + - What we need

Bayesian version of bert-baed ranking model

Bert-cqa is a bert baed ranking model… change to bayesian variant and replace gppl

Supert to warm-start for summarisation

Train at set of intervals