Text summarisation is a valuable technique that allows for the computational processing of documents, saving readers hours in manual processing. Users have different summary requirements; however, current extractive summarisation systems construct generic summaries that are not tailored to the user's needs. Asking users for feedback is one solution to combat this problem, yet this introduces an additional step of manual processing. Thus, we look to minimise the amount of required user feedback.

\medbreak

This project will investigate the feasibility of applying newly-developed techniques from Bayesian deep learning \cite{Wilson20} to get significant estimates of the model's confidence, so we can ask the user for more explanatory feedback. Legacy approaches use Bayesian optimisation \cite{Simpson19} strategies the achieve minimal user feedback; however, this strategy is blocked since modern summarisation techniques involve deep neural networks which cannot effectively express uncertainty and are typically overconfident when encountered by new topics \cite{Xu19}. This poses an issue in utilising the feedback strength of Bayesian optimisation.

\medbreak

Specifically, we look to utilise pre-trained deep learning models such as BERT to ascertain instances in a vector format to be used in an active learning component. Monte-Carlo Dropout \cite{Gal15} techniques appear to be proficient approximations for parameter posterior distributions. Thus, we will look to utilise this approach to calibrate our model.

\medbreak

It is common in passage ranking active learning solutions to use a pool-based strategy to query unlabelled instances \cite{EinDor20}. However, this requires excess computational processing. Thus, we look to use a stream-based approach to identify instances to query since it provides a lighter framework for an interactive setting. Query-by-committee acquisition functions are popular for stream-based active learning; however, since Simpson et al. \cite{Simpson19} found Bayesian optimisation strategies effectively minimised user feedback, we will look to utilise strategies such as expected improvement since Bayesian deep learning will provide a higher level of model confidence.