1. Motivation

- Learning to Rank (LTR) is the application of supervised machine learning in the construction of ranking models for information retrieval systems.
- Previous works have shown that human judges may not agree with each other in the task of assigning relevance labels to query-document pairs which induces noise in training data [1].
- Studies have shown that noise in training data for Learning to Rank negatively affects the learned ranking model [2].
- Types of Learning to Rank Algorithms: Pointwise, Pairwise, Listwise. Focus on this paper is on the pairwise algorithm which takes pairwise document preferences for each query separately to learn the ranking model. Eg. Rank SVM, Rank Net, GB Rank, QB Rank etc.
- This paper proposes a way of correcting noise in the training data for pairwise LTR algorithms.

2. Overall Objective

- **End Goal:** Show that correcting errors in training document preferences can improve Pairwise LTR performance.
- In this short paper, building upon [2], we propose a way to correct significant amount of document preferential errors automatically even if high levels of noise present in the training data.

3. Experimental Setup

- Different noise levels are injected on original training data to check efficiency and robustness of noise correction process.
- **Noise Injection:** For noise level \( p \), each pairwise document preference is reversed with probability \( p \) and kept the same with probability \( 1 - p \).
- **Noise Measurement:** Fraction of incorrect document preference pairs from the total number of preference pairs in the Post Corrected Document Preferences Set. Correctness is checked compared with Original Document Preferences Set. (Also called pNoise [2].)

4. Noise Injection and Measurement

- Studies have shown that noise in training data for Learning to Rank negatively affects the learned ranking model [2].
- This paper proposes a way of correcting noise in the training data for pairwise LTR algorithms.

5. Notation

**Partial Pairwise Document Preference Set:**
\[
\{F(q \cdot d_i > d_j) \mid \text{rel}(q, d_i) > \text{rel}(q, d_j) \text{ and } d_i, d_j \in D\}
\]

**Full Pairwise Document Preference Set:**
\[
\{F(q \cdot d_i > d_j) \cup \{F(q \cdot d_j > d_i)\mid \text{rel}(q, d_i) > \text{rel}(q, d_j) \text{ and } d_i, d_j \in D\}
\]

6. Two Phase Noise Correction Process

7. Results

8. Limitations

- **Limitations**
  - Noise Injection does not model human behaviour.
  - Efficacy of successful noise reduction on improvement of performance of Learning to Rank algorithms is not shown.

9. Acknowledgement

I acknowledge the support of ACM SIGIR in the form of a Travel Grant, which enabled me to present this poster at AIRS 2016 conference.

References
