***XAI in SLR***

***Report 2***

**How to build a good training set? –** based on ***Guidance for using artificial intelligence for title and abstract screening while conducting knowledge syntheses***

Article is about the retrospectively evaluating the implementation and performance of active- machine learning (AML) across a set of ten historical completed systematic reviews.

Methods used in the article for citation screening are available in **DistillerSR®**.

**Two approaches to using AML in systematic reviews:**

1. **Prioritization of records**:
   * Records are sorted by **the likelihood of their relevance** for further analysis.
   * While all records are screened, the most relevant ones are reviewed first. This approach:
     + **Prioritizes publications of the highest importance.**
     + Allows team members to be reallocated to other review stages
2. **Stop-Screening Rule**:
   * Screening stops once the **model reaches a predefined threshold of performance**, such as 95% estimated recall.

**Training data proccess:**

1. **Piloting**:
   * The initial set of records should be reviewed independently by at least two reviewers.
   * The size of the pilot set depends on the software:
     + **DistillerSR**: Requires screening 2% of total records (e.g., 150 records for a database of 7500).
     + **SWIFTActive-Screener**: Activates prioritization after certain conditions are met and continuously updates after every 30 records.
   * Reviewer agreement must be verified (e.g., achieving a kappa coefficient ≥ 0.8) before finalizing the pilot.
2. **Involvement of experts**:
   * Including domain experts can improve the quality of the training set, as they are better equipped to identify relevant records accurately, and accelerate early learning by the model.
3. **Targeted searches**:
   * Identifying key studies (e.g., from grant applications or previous reviews) helps refine the training set.
   * For rare diseases or niche topics, pre-selecting relevant publications aids the model in understanding the specificity of the research question.

**Practical implementation considerations:**

* **Timing of prioritization activation**:
  + In **DistillerSR**, prioritization activates after screening at least 2% of the records.
  + **Abstrackr** updates prioritization every 24 hours, requiring careful project planning.
  + **SWIFTActive-Screener** updates continuously after reviewing every 30 records.

**Key takeaway:**

How DistillerSR, Abstrackr, SWIFTActive-Screener cooperate live with the users.

Article: **Artificial intelligence for literature reviews: opportunities and challenges**

This article reviews the application of AI in SLRs, focusing on semi-automation of tasks like screening and data extraction. It provides a comprehensive analysis of tools incorporating Explainable AI (XAI) to improve transparency and usability in the citation screening process.

**Methods Used in XAI for Citation Screening**

1. **AI Algorithms**

* **Support Vector Machines (SVMs):** Most tools, like **Abstractr** and **Rayyan**, use SVM for categorizing papers as relevant or irrelevant.
* **Logistic Regression and Neural Networks:** used by **ASReview**
* **Word and Sentence Embeddings:**
  + Bag-of-Words (BoW)
  + **Colandr and Iris.ai** use **word2vec, SciBERT, and GloVe** f**or semantic understanding**.

1. **Transparency Enhancements**

* Tools document reasons for exclusion during screening to **improve reproducibility and accountability.**
* Systems **like RobotReviewer** explain decisions through supporting sentences extracted from texts.

1. **Process of Implementation**

**Initial Training Set**

* Initial set of included/excluded papers is essential.
* **Tools** **like DistillerSR and Rayyan require small seed datasets** (e.g., 1-30 papers) to start training.

**Iterative Screening Workflow**

1. **User Reviews Initial Set:**
   * Reviewers manually classify a small number of papers.
   * Data is used for AI models tools.
2. **Prioritization Activation:**
   * Once trained, the AI ranks unscreened papers by relevance.
   * Tools continuously learn from ongoing user feedback (e.g., SWIFTActive-Screener updates models every 30 reviews).
3. **Post-Screening Validation:**
   * Advanced tools like (f.e. Iris.ai) provide summaries and analyses of screened results,

**Integration of XAI Techniques**

* **Human Interaction:**
  + Interfaces allow reviewers to collaborate on the decision, to make changes in some ambiguous recommendations
* **Pre-Screening Support:**
  + Keywords and Boolean search enhance manual selection Also clustering the data makes it easy for reviewers to point out relevant studies.
* **Post-Screening Support:**
  + Summarization tools create concise outputs for review synthesis.

**Tools of XAI**

|  |  |
| --- | --- |
| **Benefits** | **Challenges** |
| Transparent workflow | Simplicity and explainability doesn’t cope great with the model accuracy |
| Decision-making efficiency | Scalability for large datasets with good transparency remaining |
| Visual interfaces making explaining easier |  |
| Updates reduce bias and errors with adapting to criteria |  |