

**Highlighter: automatic highlighting of electronic learning documents**

**Abstract:**

Electronic textual documents are among the most popular teaching content accessible through e-learning platforms. Teachers or learners with different levels of knowledge can access the platform and highlight portions of textual content which are deemed as particularly relevant. The highlighted documents can be shared with the learning community in support of oral lessons or individual learning. However, highlights are often incomplete or unsuitable for learners with different levels of knowledge. This paper addresses the problem of predicting new highlights of partly highlighted electronic learning documents. With the goal of enriching teaching content with additional features, text classification techniques are exploited to automatically analyze portions of documents enriched with manual highlights made by users with different levels of knowledge and to generate ad hoc prediction models. Then, the generated models are applied to the remaining content to suggest highlights. To improve the quality of the learning experience, learners may explore highlights generated by models tailored to different levels of knowledge. We tested the prediction system on real and benchmark documents highlighted by domain experts and we compared the performance of various classifiers in generating highlights. The achieved results demonstrated the high accuracy of the predictions and the applicability of the proposed approach to real teaching documents.