December 04, 2017

# Automatic identification and ranking of knowledge sources

Pedro Gonnet

Victor Cărbune



# Automatic identification and ranking of knowledge sources

ABSTRACT

While authoring content, users can benefit from being able to provide citations and to verify statements that include facts. This disclosure describes techniques described to provide citations and/or corrections for user text, e.g., in a document, within an authoring application. Trained machine learning models are used to generate a feature vector based on user text which is matched to feature vectors of known information sources. Citations and corrections are provided based on identifying reliable sources such as academic publications, reliable news sources, etc. along with a visual indicator of the quality of the source.

KEYWORDS

- Knowledge sources
- Word processor
- Authoring tool
- Fact verification
- Citation

## BACKGROUND

Users authoring documents, e.g., articles, papers, social network, blog posts, etc. often need to leave the authoring application (e.g., a word processing application, a blog hosting website, etc.) to find information to include in the document. For example, to include facts or cite a resource, users need to switch to another application that provides the facts.

DESCRIPTION

This disclosure describes utilization of trained machine learning model implemented within an authoring application (e.g., a word processing application) to provide citations for sources of information along with intuitive visual indicators of the quality of such sources. With user permission and express consent, user entered text or other information is automatically detected and provided to the trained model. The trained model determines and provides citations based on the user entered text. Alternatively, the user can explicitly invoke the trained model, e.g., via a request for citations for selected content.

For example, while authoring an article or blog post, users may seek citations or hyperlinks to knowledge sources that support the user-written content. For example, the user can select text that includes a statement or a fact, e.g., “Switzerland’s GDP relies heavily on the tertiary sector” or “Yeast are Eukaryotes” and request a citation that supports the statement. In response, the trained model identifies relevant citations and ranks the quality of the sources.

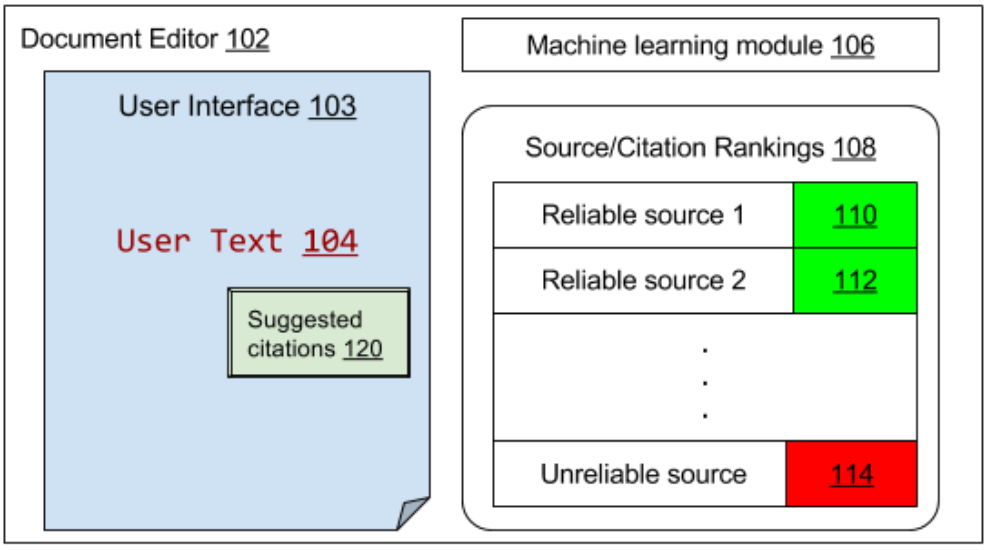


**Fig. 1 Document editor with automatic citations**

Fig. 1 illustrates a document editor (102) with automatic citations. The document editor provides a user interface (103) that enables a user to enter text, e.g., user text (104). The document editor includes a machine learning module (106) that ranks citations or sources (108) corresponding to user text and provides the suggested citations (120).

With user permission, the machine learning module that includes a trained model, can scan text or other information entered by the user and automatically provide citations. Alternatively, the user can explicitly request citation for a particular portion of the document. As illustrated in Fig. 1, sources 110, 112, and 114 are identified by a trained machine learning model based on the user text. The sources are ranked based on their reliability. For example, sources such as peer-reviewed journal or conference articles, encyclopedias, reputed newspapers and magazines, etc. are ranked high (indicated with green color) and other sources that are unknown or known to be unreliable, e.g., unranked blogs, etc. are ranked low (indicated with red color). Further, when the user text has inaccuracies, the model can suggest corrections based on a reliable source, via a visual indicator and a link to the source.

Further, the model can also search for the inverse of user text, when the user enters a factual statement. Such a search can provide a ratio of sources that support the user text to sources that indicate that user text is inaccurate, providing additional data to the user. Further, other known scores, e.g., ranking of web pages, etc. can also be incorporated.

The techniques can also be deployed in contexts other than document authoring. For example, upon user request, the techniques can also analyze incoming content (e.g., messages) from a third-party and provide visual indicators (e.g., color coding) of the quality of such content along with relevant citations. The techniques can also be used for other types of content, e.g., images. For example, the techniques can be used to rank images based on their credibility. For

example, an image that has been manipulated or artificially generated can be ranked lower compared to the original.

The techniques described can use a machine learning model to convert user text, e.g., a single statement, into a feature vector and knowledge sources that include a set of similar feature vectors, e.g., generated using a different machine learning model. The goal of the second model is to learn to map parts of the referenced source to the text referencing the source. During a citation or reference search, the first model is used to convert the selected statement into a feature vector, which is then used together with the second model to match it with similar parts of possible reference sources from a known database of sources, that might be stored or provided as a result of processing a query using an external database system. Both models can be trained with publicly available citations and information sources.

The training data can include academic research papers, e.g., with every statement in a paper comprising a data point. Other training data can include news articles, blogs, etc. that include hyperlinks to other articles. These data sources can be used to train both models simultaneously by extracting feature vectors for a statement and the source, and matching the feature vectors as described above. During training, statements and sources can be paired such that each statement feature vector matches one or more of the source feature vectors, but not any of the other unrelated source feature vectors. Expert systems can be used instead of a machine learning model to identify matching sources. However, such implementation may be more expensive, e.g., due to a need to process each information source.

The described techniques can be implemented in various software applications, e.g., document authoring tools such as word processors, email clients, virtual assistants, browsers, etc.

In situations in which certain implementations discussed herein may collect or use personal information about users (e.g., user data, information about a user's social network, user's location and time at the location, user's biometric information, user's activities and demographic information), users are provided with one or more opportunities to control whether information is collected, whether the personal information is stored, whether the personal information is used, and how the information is collected about the user, stored and used. That is, the techniques discussed herein collect, store and/or use user personal information specifically upon receiving explicit authorization from the relevant users to do so.

For example, a user is provided with control over whether programs or features collect user information about that particular user or other users relevant to the program or feature. Each user for which personal information is to be collected is presented with one or more options to allow control over the information collection relevant to that user, to provide permission or authorization as to whether the information is collected and as to which portions of the information are to be collected. For example, users can be provided with one or more such control options over a communication network. In addition, certain data may be treated in one or more ways before it is stored or used so that personally identifiable information is removed. As one example, a user's identity may be treated so that no personally identifiable information can be determined. As another example, a user's geographic location may be generalized to a larger region so that the user's particular location cannot be determined.

## CONCLUSION

While authoring content, users can benefit from being able to provide citations and to verify statements that include facts. This disclosure describes techniques described to provide citations and/or corrections for user text, e.g., in a document, within an authoring application.

Trained machine learning models are used to generate a feature vector based on user text which is matched to feature vectors of known information sources. Citations and corrections are provided based on identifying reliable sources such as academic publications, reliable news sources, etc. along with a visual indicator of the quality of the source.