August 20, 2018

SYSTEM AND METHOD FOR REMOVING UNWANTED CONTENT FROM GETTING PRINTED USING MACHINE LEARNNG TECHNIQUES

HP INC

System and Method for removing unwanted content from getting printed using machine learning techniques

Parental control for printers

Abstract:

Printers are used to print content for purposes of sharing, and therefore, it becomes necessary to control and monitor the content that is being printed. At home scenario, it could be that the parents would like to control what their kids print, whereas at enterprises, it becomes necessary to ensure forbidden content is not allowed to print. The current document outlines a way by which this can be achieved. Machine learning is used to identify patterns within the content that are not allowed, and thereby avoid them getting printed. The approach uses a content packetizer to packetize the stream, and then a classifier to decide if the content should be printed or not.

Problem Statement

Enterprises and home users alike would like to control the content that gets printed on their printer to ensure that it is within "permitted" range. By this, it is meant that the content is within allowed legal bounds, does not contain objectionable content and does not contain material that is harmful or dangerous to the user. In the home scenario, parents might want to control the printing that their kids do to allow only allowed content. Enterprises might want to ensure that certain content that is not allowed (for ex, confidential legal content, etc) is not printed by its users. In the current disclosure, a method is provided to achieve the same.

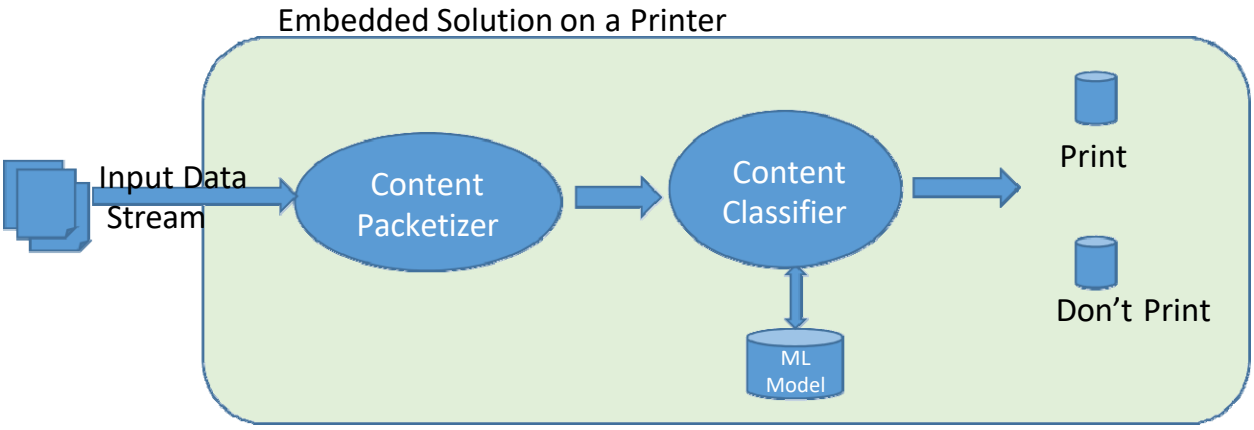
The idea

The idea is to create a classifier based on machine learning which processes the data coming to the printer, and provides a label (with a confidence indicator) about the suitability of the content for printing. This classifier could run on the device or off the device. It enables a user to select the class of the content that can be printed on a printer. It allows a user to select the allowed content (by using either sample pages of disallowed text, or a set of disallowed words, or a set of disallowed images). It then, looks for similarity of the print content with the specified content, and only if the content is different from the pre-specified barred content, it is allowed to print. Otherwise, the content is discarded.

We outline the option with the disallowed words in this disclosure.

Description:

Please refer to the diagram given below:



The steps are given below:

1. The input data stream is received by the device (printer/scanner) over the network/USB
2. The content packetizer processes the data stream, and breaks up the stream into individual print documents. In other words, the content packetizer discretizes the print content. The packetizer looks for "document start" and "document end", and uses these markers to discretize the content. The data from "document start" to "document end" becomes the print document.
3. The content classifier is a classifier that is trained using machine learning algorithms. The classifier reads a pre-trained ML model, and when a certain incoming document is received, it classifies as print-ready or print-not-ready. The classifier classifies the document as print-ready, if the document does not contain any thing in content that is found (or similar to) disallowed content.
4. The documents that are labelled as print-ready, are forwarded to the print subsystem for printing. The documents that are labelled as "print-not-ready" are logged, and discarded. They are not printed.

The content classifier uses multiple approaches for document classification. The below computation is for only words (that too. disallowed ones), a similar approach is used for images as well. All the content is treated as vectors (for example, word vectors) so that the classification process is reasonably fast.

1. Initially it loads the list of disallowed-allowed content, and searches for occurrence of this content in the incoming document. It computes the probability number based on this pattern match.
   disallowed content list = {list of all prohibited content}
   The unit of comparison depends on the provided disallowed content – for example, for text content we use words, sentences or documents, and for images, we can use images itself.

   If n = number of disallowed words that are found in the incoming document, and
   N = total number of words in the incoming document, then
   the (Ppw) probability of prohibited words = n/N

   If Ppw < threshold, then the document is printed
   If Ppw >= threshold, then the document is discarded

The threshold is set by the user. We have shown the computation for words, but a similar algorithm computes the probability for images as well.

Training:

During training phase, the classifier is trained with content that is rich in disallowed content. The corpus for training is created manually with text and images from disallowed content set. At the end of the training, the model is trained to identify any occurrence of word or image in the provided content that is not allowed to print.

Testing/Production:

During the testing phase, we provide some sample content which is different from the training samples, to validate the model that was created earlier. We then compute the accuracy of the model (both the precision and recall). This helps us to tune the model if necessary by providing additional training content.

Advantages

1. Printing only content that is permitted. Uses a well known metaphor (Parental control). This could be an app or a button.

Appendix A:
One possible way to enable/disable parental control is to use a button on the control panel.

Appendix B:
There are different ways to build and deploy the solution. Once possible option is to use an IO Filter solution.

Disclosed by Raghu Anantharangachar, HP Inc.