Webthetics: Quantifying Webpage Aesthetics with Deep Learning

Avery Anjard

Internet Programming

Professor Jim Littleton

July 23rd, 2023

The aesthetic value of a webpage can often be a difficult thing to quantify using scientific methods. Creating an aesthetically pleasing design that draws users in is often considered an art form that can only be measured subjectively. However, Dou et al. have created a deep-learning model, which they call Webthetics, that is trained on user ratings of webpages and can quite accurately predict how a user will rate the aesthetic of a new webpage.

The researchers used a type of deep-learning network called a Convolutional Neural Network, or CNN [1]. The dataset used to train their model is a set of 398 webpages that have been rated by 40,000 users based on their aesthetic value. A Likert scale from 1 to 9 was used to rate each webpage. It was important that during the training of the model, the bigger the difference between the predicted score and the actual score, the bigger the loss should be. That is, the model should be “punished” more for being further away from the actual score. For this reason, the researchers used a regression-based loss, which gives predictions on a continuous scale, instead of a classification-based one, where the predictions are discrete integers.

The dataset was divided into a training set with 300 webpages and a testing set with 98 [1]. Given that the dataset is small, the model may overfit the data. To prevent overfitting in their model, they used a method called transfer learning, where they take a preexisting model – in this case, a model that was trained on an image prediction dataset – and use it as the base for their model. In their model, the first layers are based on more generalized features, that can be applied to any image dataset, whereas the later layers become more specialized for predicted webpage ratings.

To measure the accuracy of their model, the researchers used the Pearson coefficient [1]. This value goes from -1 to 1, where -1 shows a perfect negative correlation, and 1 indicates a perfect positive correlation. In other words, the higher the value, the more accurate the model is. The results of the model’s tests show that it is quite accurate, with a Pearson coefficient of 0.85. In the dataset, users commonly rated webpages from 4 to 6, with 62% of ratings being in this range; the model accurately predicted this, with 63% of its predictions being in this range.

Dou et al. also compared their deep-learning model to other kinds of models. For instance, they compared a model with no transfer learning (i.e., they did not use the base model), a model with soft transfer (i.e., the first two layers from the image model were used), and a model with hard transfer (i.e., all the layers from the mage base model were used) [1]. They found that the no-transfer model was significantly less accurate than both the soft and hard transfer models, which were relatively similar. The no-transfer model had a Pearson coefficient of only 0.56, while the soft and hard transfer models had coefficients of 0.85 and 0.83, respectively. The researchers also compared their model to a handcrafted one, where the model measures the colorfulness and complexity of the webpage and uses those two metrics to predict a user’s aesthetic rating. The Pearson coefficient of this model was only 0.59, much less than the researchers’ deep learning model.

For their final tests, Dou et al. measured their model’s sensitivity to image manipulation by distorting a test webpage and seeing how the model’s accuracy changed [1]. They first put 100 small squares randomly over the webpage, each being 10 by 10 pixels big. They then increased the size by 10 pixels on each side several times, up to 50 by 50. The Pearson coefficient decreased rapidly each time, showing that the model is quite sensitive to the layout structure of a webpage. The researchers conducted similar tests by using one larger black square form size 100 by 100 pixels up to 500 by 500, occluding the webpage with different percentages from left to right, and putting the webpage through both a high pass filter and a low pass filter. They were able to conclude that their model is somewhat sensitive to content contrast and organizational balance (though not to as high of a degree as the other measurements).

Dou et al. conclude that they have created an effective new way of quantifying the aesthetic value of a web page with their deep-learning model [1]. They show that their model is more accurate than previous handcrafted models, as well as other deep-learning models, which did not use regression modeling or transfer learning. They also note that since aesthetic values tend to change over time, their deep-learning model will automatically update over time, whereas the handcrafted models would have to be changed manually.

This new model will likely have a far-reaching impact on the web-design world. Creating a website that will draw in users and keep them can be a difficult thing to do, considering that aesthetic values can be so subjective. However, there is still a general idea of what an aesthetically appealing website looks like for most people, and Dou et al.’s model will make creating such a website much simpler.

[1] Q. Dou, X. S. Zheng, T. Sun, and P.-A. Heng, “Webthetics: Quantifying webpage aesthetics with Deep Learning,” *International Journal of Human-Computer Studies*, vol. 124, pp. 56–66, Apr. 2019. doi:10.1016/j.ijhcs.2018.11.006