Multimodal Side-Tuning for Document Classification

25th International Conference on Pattern Recognition

Accepted paper N. 1425

Stefano P. Zingaro, Giuseppe Lisanti, and Maurizio Gabbrielli
Department of Computer Science and Engineering (DISI)
University of Bologna, Italy
Automatic Document Image Classification
combining and exploiting

- Visual + textual features (multimodal input sources)
- Lightweight deep Neural Networks (e.g. MobileNetV2, ResNet50)
- Network adaptation by Side-Tuning (Zhang et al., 2020)
RVL-CDIP document image dataset

400,000 images from 16 categories

- letter
- memo
- email
- filefolder
- form
- handwritten
- invoice
- advertisement
- budget
- news article
- presentation
- scientific publication
- questionnaire
- resume
- scientific report
- specification

Image from cs.cmu.edu/~aharley/rvl-cdip.
THE NICOTINE'S WHAT MAKES 'EM SMOKER!

By W. K. Collins

An in-depth look at the primary ingredient that affects your business.

The stimulation from nicotine in tobacco is the main reason 60 million Americans smoke.

Numerous other chemicals are present in tobacco smoke, and it is the interrelationship of these constituents that provides the total pleasure derived from smoking. But nicotine is component which provides the physiological stimulation that makes a person a compulsive smoker.

Fortunately, the Surgeon General's report on tobacco and health gave nicotine a clean bill of health.

Tobacco manufacturers regulate the level of nicotine in their products very closely. Costs of tobacco going into a product and nicotine level of the final product receive top attention as cigarettes are raised. Over a period of time the percentage of nicotine in products may vary, but abrupt changes are avoided. Usually it takes a 10% to 25% change for a significant part of the public to detect it. Nevertheless, it's taboo within the industry to make unnecessary changes.

The nicotine level in most best selling cigarettes is currently somewhere between 1.5% and 2%. Some cigarettes are available with much less nicotine, with 25% of the market in 1977 considered low nicotine. High-nicotine cigarettes fail to attract and keep many smokers. The lack of nicotine in so-called no-nicotine cigarettes has surely played a major role in the downfall of these products.

There's a definite correlation between smokeable nicotine in tobacco and its aroma or pleasing qualities during combustion. High-nicotine tobaccos are too strong; low-nicotine cigarettes are too weak. To this country they found the Indians smoking high-nicotine (about 10%) tobaccos. The high-nicotine tobaccos were simply too strong for our people.

"Bleeding" is the main method used by manufacturers to regulate the desired nicotine levels in tobacco products. The basic blend of our cigarettes is composed of four classes of tobacco, each of which has distinct properties. In 1976 the average percentage of each class of tobacco used in cigarettes was: Burley, 46%; Maryland, 21%; and imported, 18%.

Nicotine contributed by the high-cured portion of the blend is controlled by using tobacco grown in different years, from different production areas of the Belt, and from different stalk positions. Nicotine content of tobacco grown in different years and production areas varies considerably. It's mainly related to rainfall.

Most of the 60 species of tobacco are very low in nicotine. (We'll define a species as a group of freely interbreeding plants.) Nicotine can be artificially produced, but to date this process has been cost prohibitive for commercial use. However, efforts by private companies are currently underway to make synthetic cigarettes with nicotine added.

Buying patterns of companies are influenced by the levels of nicotine of tobacco in inventory as well as by what is being purchased. As leaf inventories held by companies...
THE NICOTINE’S WHAT MAKES ‘EM SHOE!
By &. K. Collins
An in-depth look at the primary ingredient that affects your business.

The stimulation from nicotine in tobacco is the main reason that 60 million Americans smoke. Numerous other chemicals are present in tobacco smoke, and it is the interrelationship of these constituents that provides the total pleasure derived from smoking. But nicotine is the major component which provides the physiological stimulation that makes a person a compulsive smoker.

Recently, the Surgeon General’s report on tobacco and health gave nicotine a clean bill of health.

Tobacco manufacturers regulate the level of nicotine in their products very closely. Costs of tobacco going into a product and nicotine level of the final product receive top attention as cigarettes are made. Over a period of time the percentage of nicotine in Products may vary, but abrupt changes are avoided. Usually it takes a 15% to 20% change for a significant part of the public to detect it. Nevertheless, it’s taboo within the industry to make unnecessary changes. The nicotine level in most big selling cigarettes is currently some 8.5%. Some cigarettes are available with much less nicotine, with 25% of the market in 1977 considered low nicotine. High-nicotine cigarettes also fail to attract and keep many smokers.

The lack of nicotine in so-called no-nicotine cigarettes has surely played a major role in the downfall of these products.

There’s a definite correlation between the level of nicotine in tobacco and its aroma or pleasurable qualities during combustion. Tobacco tobaccos usually are weak flavored.

High-nicotine tobaccos are too strong for most American smokers. When our forefathers to this country they found the Indians smoking high-nicotine (about 10%) tobacco, The tobaccos were simply too strong for our people. Blending is the main method used by manufacturers in tobacco products. The basic blend of our cigarettes is composed of four classes of tobacco, each of which has distinct properties. In 1976 the average percentage of each class of tobacco used in cigarettes was: flue-cured, 46%, Burley, 34.7, Maryland, 22; and imported, 18%.

Nicotine contributed by the flue-cured portion of the blend is controlled by using tobacco grown in different years, from different production areas of the Belt, and from different Stalk positions. Nicotine content of tobacco grown in different years and production areas varies considerably. It’s mainly related to rainfall.

Most of the 60 species of tobacco are very low in nicotine, (We’ll define a species as a group of freely interbreeding plants.) Nicotine can be artificially produced, but to date this process has been cost prohibitive for commercial use. However, efforts by private companies are currently underway to make synthetic cigarettes with nicotine added.

Buy inventory as well as of what is being purchased. But as leaf inventories held by companies 576106781
Network Adaptation via Additive Side Network

sidetuning.berkeley.edu (Zhang et al., 2020)

\[ \alpha \cdot \text{Locked} + (1-\alpha) \cdot \text{Unlocked} \]
Multimodal Side-Tuning

\[ \alpha \cdot \text{Locked} + \beta \cdot \text{Unlocked} + 1-(\alpha+\beta) \cdot \text{Locked} \]

\[ \ldots \]

Tesseract OCR
Alpha Curriculum
Side-Tuning to explore Document Classification models

$\alpha$

$\beta$

0,5

0

1

0,5

1

IMAGE

FINE-TUNING

TEXT FROM SCRATCH

IMAGE FEATURE EXTRACTION

$\alpha$

$\beta$
## Prediction performance improvement

Sample of per-class accuracy achieved on Tobacco3482 dataset

<table>
<thead>
<tr>
<th>Method</th>
<th>Adve</th>
<th>Email</th>
<th>Form</th>
<th>Letter</th>
<th>Memo</th>
</tr>
</thead>
<tbody>
<tr>
<td>CNN for text classification from scratch</td>
<td>93%</td>
<td>30%</td>
<td>77%</td>
<td>59%</td>
<td>50%</td>
</tr>
<tr>
<td>Image fine-tuning with MobileNetV2</td>
<td>97%</td>
<td>97%</td>
<td>83%</td>
<td>84%</td>
<td>87%</td>
</tr>
<tr>
<td>Multimodal fine-tuning (Audebert et al., 2020)</td>
<td>93%</td>
<td>98%</td>
<td>88%</td>
<td>86%</td>
<td>90%</td>
</tr>
<tr>
<td>Image side-tuning with MobileNetV2</td>
<td>89%</td>
<td>97%</td>
<td>87%</td>
<td>90%</td>
<td>96%</td>
</tr>
<tr>
<td>Multimodal side-tuning with MobileNetV2</td>
<td>95%</td>
<td>99%</td>
<td>89%</td>
<td>93%</td>
<td>93%</td>
</tr>
<tr>
<td>Multimodal side-tuning with ResNet50</td>
<td>96%</td>
<td>98%</td>
<td>91%</td>
<td>92%</td>
<td>93%</td>
</tr>
</tbody>
</table>
Lightweight architecture

Overall Accuracy on RVL-CDIP dataset w.r.t. #Parameters
Grazie!

szingaro.github.io/multimodal-side-tuning