

Improving Google Colaboratory to serve Thailand Machine Learning Community

Korakot Chaovavanich

VISTEC-depa Thailand Artificial Intelligence Research Institute

Colab, or Google Colaboratory, was released in October 2017. It has now become a popular tool for knowledge sharing in the machine learning community. Although partly open sourced, its development is strictly controlled by Google, and does not accept external contributions. To make Colab serve Thai community better, an external library package is needed. This paper presents the architecture of such a library: kora.

The objective of this library is to add features to Colab from community requests; a way to add Thai fonts, for example. After 5 months of development, the structure of the library reflects various needs of the community:

1. Installing software packages easily (e.g. `kora.install.mysql8`)
2. Downloading Thai datasets easily (e.g. `kora.datasets.ORCHID`)
3. Additional data visualization tools (e.g. `kora.viz.wordcloud`)
4. Experimenting with latest AI research (e.g. `kora.ai.colorizer`)

The library now has more than 50 features. Notable features are:

- `docker`: installing software from a docker image
- `jupyterlab`: using traditional notebook and lab interface, allowing Python 3.8
- `kaggle`: downloading datasets directly from kaggle
- `console`: to run Linux commands easily
- `ngrok`: for running web application software

The library is still new and not yet widely known in Thailand machine learning community. Its impact is expected to grow alongside Google Colab. Currently, the library is used in the SuperAI Engineer program. It is distributed through PyPI and has now more than 1,000 downloads per week.

The source code is available from <https://github.com/airesearch-in-th/kora>.