

# Designing a Visual Analytics System for Industry-Scale Deep Neural Network Models

Minsuk Kahng\*  
Georgia Tech

Pierre Y. Andrews†  
Facebook

Aditya Kalro‡  
Facebook

Duen Horng (Polo) Chau§  
Georgia Tech

## ABSTRACT

The complexity of industry-scale deep learning models and datasets pose unique design, visualization, and system challenges that are inadequately addressed by existing work. Through participatory design sessions with over 15 researchers and engineers at Facebook, we have designed, developed, and deployed **ACTIVIS**, a visual analytics system for interpreting industry-scale deep learning models and results. By tightly integrating multiple coordinated views, such as a computation graph overview of the model architecture, and a neuron activation view for pattern discovery and comparison, users can explore complex deep neural network models at both instance- and subset-level. **ACTIVIS** has been deployed on Facebook’s machine learning platform.

This article is a summary for the VAST’17 paper (TVCG track) *ActiVis: Visual Exploration of Industry-Scale Deep Neural Network Models* [2].

## 1 DESIGNING FOR INDUSTRY-SCALE MODELS

Despite the increasing interest in developing visualization tools for deep learning interpretation [5–7], the complexity of large-scale models and datasets used in industry pose unique design challenges that are inadequately addressed by existing work. For example, while most existing visualization tools target image datasets, deep learning tasks in industry often involve different types of data, including text and numerical data. Furthermore, in designing tools for real-world deployment, it is a high priority that the tools be flexible and scalable, adapting to the wide variety of models and datasets used. These observations motivate us to design and develop **ACTIVIS** [2], a visual analytics system for deep neural network models, now deployed on Facebook’s machine learning platform.

Since the **ACTIVIS** project started in April 2016, we have conducted participatory design sessions with over 15 Facebook engineers, researchers, and data scientists across multiple teams to learn about their visual analytics needs. We identified six key design challenges — for *data*, *model*, and *analytics* — that have not been adequately addressed by existing deep learning visualization tools. The challenges include the need to support: (1) diverse input data sources, (2) high data volume, (3) complex model architecture, (4) a great variety of models, (5) diverse subset definitions for analytics, and (6) both instance- and subset-level analyses. These challenges shape the main design goals of **ACTIVIS**.

## 2 ACTIVIS CONTRIBUTIONS

**ACTIVIS**’s main contributions include:

- A novel visual representation that unifies instance- and subset-level inspections of neuron activation, facilitating comparison of activation patterns for multiple instances.

---

\*e-mail: kahng@gatech.edu

†e-mail: mortimer@fb.com

‡e-mail: adityakalro@fb.com

§e-mail: polo@gatech.edu

- An interface that tightly integrates an overview of graph-structured complex models and local inspection of neuron activations, allowing users to explore the model at different levels of abstraction.
- A deployed system scaling to large datasets and models.
- Case studies with Facebook engineers and data scientists that highlight how **ACTIVIS** helps them with their work.

**ACTIVIS**’s multiple coordinated views help users get a high-level overview of the model from which the user can drill down to perform localized inspection of activations. **ACTIVIS** visualizes how *neurons* are activated by user-specified instances or instance subsets, to help users understand how a model derives its predictions. The subsets can be flexibly defined using data attributes, features, or output results, enabling model inspection from multiple angles. While many existing deep learning visualization tools support instance-level exploration [6, 7], **ACTIVIS** is the first tool that simultaneously supports instance- and subset-level exploration. Both exploration strategies are common and effective, and they offer complementary analytics benefits. Instance-based analysis instructs how individual instances contribute to a model’s accuracy, but it is tedious to inspect many instances one by one. Subset-based analysis leverages input features or instance subsets to help reveal relationships between data attributes and machine learning algorithms’ outputs [3]. It is especially beneficial when dealing with huge datasets in industry, which may consist of millions or billions of data points. By exploring instance subsets and enabling their comparison with individual instances, users can learn how their models respond to many different slices of the data.

We refer our readers to the longer version of our **ACTIVIS** [2] VAST’17 paper published in *IEEE Transactions on Visualization and Computer Graphics*.

## Deployment

We have deployed **ACTIVIS** on **FBLearn**, Facebook’s machine learning platform [1]. **ACTIVIS** can support as many different kinds of data types and models as what **FBLearn** currently does. Developers who want to use **ACTIVIS** for their model can easily do so by adding only a few lines of code, which instructs their models’ training process to generate information needed for **ACTIVIS**’s visualization. **ACTIVIS** users at Facebook (e.g., data scientists) can then train models and use **ACTIVIS** via *FBLearn Flow* [1], Facebook’s internal machine learning web interface, without writing any additional code.

## 3 ILLUSTRATIVE SCENARIO

To illustrate how **ACTIVIS** works in practice, consider our user Susan who is training a word-level *convolutional neural network* (CNN) model [4] to classify question sentences into one of six categories (e.g., whether a question asks about *numeric* values, as in “*what is the diameter of a golf ball?*”). Her dataset is part of the TREC question answering data collections<sup>1</sup>.

Susan starts model training with default training parameters. After training completes, she launches **ACTIVIS**, running in a web browser. **ACTIVIS** displays a model architecture through a computation graph

---

<sup>1</sup><http://cogcomp.cs.illinois.edu/Data/QA/QC/>

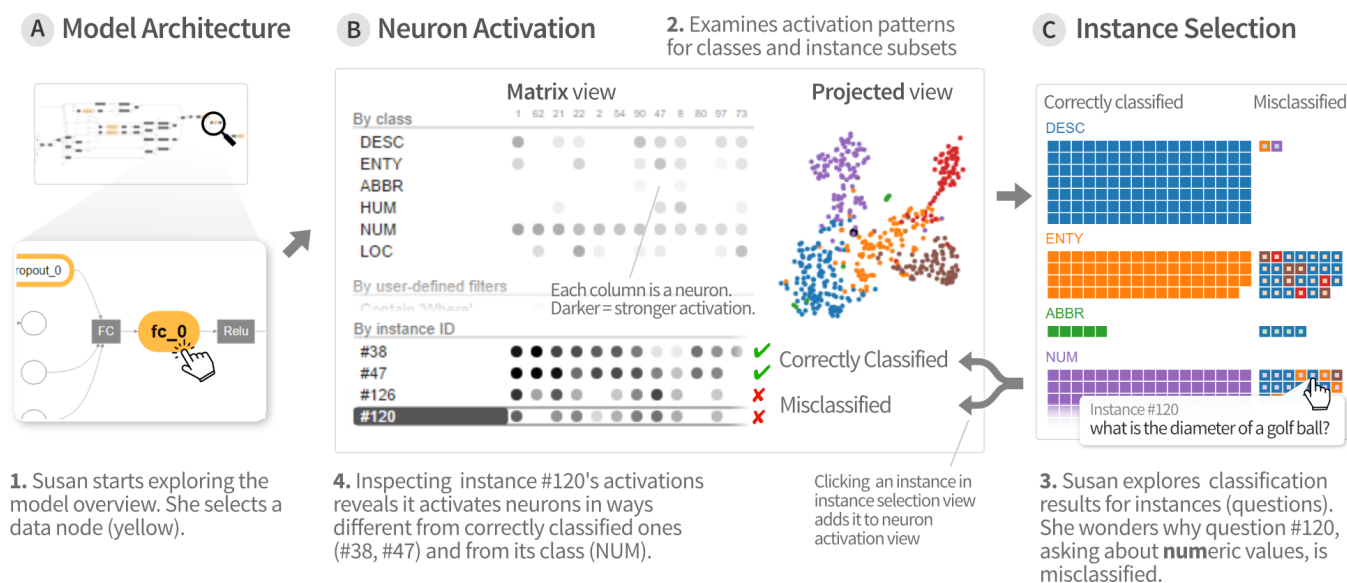


Figure 1: ACTIVIS’s multiple coordinated views supports exploration of complex deep neural network models, at both instance- and subset-level. 1. Our user Susan starts exploring the model architecture, through its *computation graph* overview (at A). Selecting a *data node* (in yellow) displays its *neuron activations* (at B). 2. The *neuron activation matrix view* shows the activations for instances and instance subsets; the *projected view* displays the 2-D projection of instance activations. 3. From the *instance selection* panel (at C), she explores individual instances and their classification results. 4. Adding instances to the matrix view enables comparison of activation patterns across instances, subsets, and classes, revealing causes for misclassification.

overview (Fig. 1A, top). By exploring the graph, Susan learns about how operations (e.g., convolution) are performed in the model.

Based on her experience working with other deep learning models, Susan knows that a model’s performance is strongly correlated with its last hidden layer, thus it would be informative to analyze that layer. So she clicks the node for the last hidden layer (highlighted in yellow, in Fig. 1A, bottom), then ACTIVIS displays the layer’s *neuron activation* in a panel (Fig. 1B): the *neuron activation matrix view* on the left shows how neurons (shown as columns) respond to instances from different classes (rows); and the *projected view* on the right shows the t-SNE projection of instance activations. From the *matrix view* where stronger neuron activations are shown in darker gray, Susan sees that the activation patterns for the six classes (rows) are quite visually distinctive, which may indicate satisfactory classification. However, in the *projected view*, instances from different classes are not clearly separated, which suggests some degree of misclassification.

To examine the misclassified instances and to investigate why they are mislabeled, Susan brings up the *instance selection panel* (Fig. 1C). The classification results for the **NUMBER** class alarms Susan, as many instances in that class are misclassified (shown in right column). She wants to compare the activation patterns of the correctly classified instances with those of the misclassified. So she adds two correct instances (#38, #47) and two misclassified instances (#120, #126) to the *neuron activation view* — indeed, their activation patterns are very different (Fig. 1.4).

Taking a closer look at the *instance selection panel*, Susan sees that many instances have **blue** borders, meaning they are misclassified as **DESCRIPTION**. Inspecting the instances’ text reveals that they often begin with “What is”, which is typical for questions asking for descriptions, though they are also common for other question types. To understand the extent to which instances starting with “What is” are generally misclassified by the model, Susan creates an *instance subset* for them, and ACTIVIS adds this subset as a new row in the *neuron activation view*. Susan cannot discern any visual patterns

from the subset’s seemingly scattered, random neuron activations, suggesting that the model may not yet have learned effective ways to distinguish between the different intents of “What is” questions. Based on this finding, she proceeds to train more models with different parameters (e.g., consider longer *n*-grams) to better classify these questions.

## ACKNOWLEDGMENTS

We thank Facebook Applied Machine Learning Group, especially Yangqing Jia, Andrew Tulloch, Liang Xiong, and Zhao Tan for their advice and feedback. This work is partly supported by the NSF Graduate Research Fellowship Program (DGE-1650044).

## REFERENCES

- [1] P. Andrews, A. Kalro, H. Mehanna, and A. Sidorov. Productionizing machine learning pipelines at scale. In *ML Systems Workshop at ICML*, 2016.
- [2] M. Kahng, P. Y. Andrews, A. Kalro, and D. H. Chau. ActiVis: Visual exploration of industry-scale deep neural network models. *IEEE Transactions on Visualization and Computer Graphics*, 24(1), 2018.
- [3] M. Kahng, D. Fang, and D. H. Chau. Visual exploration of machine learning results using data cube analysis. In *Workshop on Human-In-the-Loop Data Analytics at SIGMOD*. ACM, 2016.
- [4] Y. Kim. Convolutional neural networks for sentence classification. In *EMNLP*, 2014.
- [5] M. Liu, J. Shi, Z. Li, C. Li, J. Zhu, and S. Liu. Towards better analysis of deep convolutional neural networks. *IEEE Transactions on Visualization and Computer Graphics*, 23(1), 2017.
- [6] D. Smilkov, N. Thorat, C. Nicholson, E. Reif, F. B. Viégas, and M. Wattenberg. Embedding Projector: Interactive visualization and interpretation of embeddings. In *Workshop on Interpretable Machine Learning in Complex Systems at NIPS*, 2016.
- [7] J. Yosinski, J. Clune, A. Nguyen, T. Fuchs, and H. Lipson. Understanding neural networks through deep visualization. In *Workshop on Visualization for Deep Learning at ICML*, 2016.