**DeepBase: Scalable Inspection of Deep Neural Networks**

**Thibault Sellam, Kevin Lin, Ian Huang, Yiliang Shi, Yiru Chen**
**Carl Vondrick, Eugene Wu**

**Computer Science**
**Columbia University**

**Background — Deep Neural Inspection**

Neural networks (NNs) are revolutionizing a wide range of machine intelligence tasks with impressive performance. A rapidly growing ecosystem of development tools have made them popular and accessible.

**Major challenge:** understanding their internal logic and ensuring that they behave reliably.

Popular approach: run the model on test data and analyze the activation of the hidden units.

---

**Problem — Many Prototypes, no API**

ML engineers must implement their own interpretability tools, because:

- Many methods have little to no public implementation
- Most existing implementations are ad hoc: framework-specific and/or model-specific
- Few implementations are optimized

Result: a sparse collection of task-specific prototypes with no common API.

---

**Our System — Deep Base**

DeepBase executes and optimizes Deep Neural Inspection queries over a given collection of models, data and hypotheses.

---

**Architecture**

DeepBase queries are compiled into workflow of behavior extractors, processors and statistical scoring aggregators (called *Inspectors*).

---

**Optimizations**

We develop optimizations based on GPU parallelism, streaming, sampling, and model merging.

---

**Next Steps**

Build and maintain libraries of hypotheses based on recent ML findings

Scale the system up — better support for multicores and shared-nothing clusters

Applications in NLP, computer vision, fairness and social sciences.