



NVIDIA-powered Data Science Workstation: Taxi Fare Prediction

Enabling big data analytics at the speed of thought with OmniSci and RAPIDS

OVERVIEW

One of the core use cases for the data science workstation is exploratory big data analytics -- being able to load, prepare, visualize, and extract insights from massive datasets as fast as a given mission requires. OmniSci's GPU-accelerated analytics platform combines on-GPU databases, the RAPIDS data science platform, and strong visualization capabilities to help customers understand their data.

The demo shows an application of GPU-accelerated analytics: fare prediction on the NYC Taxi dataset in which the model was able to achieve an accuracy of under \$2. The demo also shows how you gain insights into tip amounts, as well as where and when to best deploy larger capacity vehicles.

As the customer develops an understanding of the datasets, the same OmniSci capabilities can be employed at data center scale to enable production workflows, dashboard visualizations, and other mission systems.



KEY POINTS

1. GPU-accelerated databases, the RAPIDS data science platform, and GPU-driven visualizations combine to enable interactive big data analytics on a single workstation.
2. NVIDIA Quadro RTX GPUs have the memory capacity to run sophisticated machine learning algorithms on large mission-relevant datasets (hundreds of millions or billions of rows).
3. The OmniSci software which combines these capabilities is being applied in substantial use-cases across [multiple industries](#), including [the federal government](#).