

Soldier Retention with SambaNova

World record performance and accuracy for secure generative AI

Challenge

The United States military is facing a decrease in retention rates. In 2021, the retention rate was on 55%. This has been attributed to a number of issues such as inadequate compensation, toxic leadership, and limited personal freedom.

Objective

Utilize SambaNova predictive analytics to analyze factors contributing to employee turnover, such as job satisfaction or command climate surveys, performance metrics, and external data sources. Then use the results to develop AI-driven strategies to enhance job satisfaction, customize career paths, and create a positive work environment, ultimately improving retention rates.

Distinguishing Factors

- Domain tuned expert models
- World record performance and accuracy
- Accuracy testing tools
- Data security & privacy
- Model/data ownership
- Full stack solution
- On-site/cloud services

Justification

The SambaNova platform can use predictive analytics can identify factors contributing to employee turnover with high accuracy. SambaNova Suite's ability to "match or surpass state-of-the-art closed-source models on EGAI benchmarks" ensures that predictions are reliable, helping to develop effective strategies to improve job satisfaction and retention rates.

Solution

SambaNova Suite, combined with the Samba-1 Composition of Experts (CoE) model, delivers the highest combination of performance and accuracy for generative AI. A complete hardware/software platform, SambaNova Suite can be deployed as a fully configured rack-level solution either on-premises, including air-gapped environments, or as a cloud-based solution. The Samba-1 model provides state-of-the-art accuracy across a wide range of use cases.

Key Features

Performance, Accuracy, and Efficiency

Only SambaNova delivers performance of over 1000 tokens/s at full precision, on as few as 16 sockets. This is an unrivaled combination of performance and accuracy with a small footprint, dramatically reducing power consumption.

Total AI Stack Control

With options for on-premises and air-gapped deployments, SambaNova ensures complete data privacy and security, enabling the use of sensitive data for model training without external exposure.

Model Ownership

Once a model is fine-tuned, it becomes the property of the customer in perpetuity, eliminating model ownership concerns.

Configurable Access and Permissions

The Samba-1 model enables flexible access controls, allowing precise management of model permissions to align with organizational roles, enhancing security and ensuring compliance.

SambaNova Suite delivers the most accurate, integrated full stack generative AI platform, optimized for enterprise and government organizations. Delivered through SambaNova's integrated AI platform, which can be deployed on-premises or through the cloud, SambaNova Suite seamlessly integrates into existing business processes to deliver transformative capabilities. With the capacity to be further refined using customer data, these models can deliver unparalleled accuracy while providing enterprise grade security and data governance.

Learn more

