Optimizing Natural Language Processing ML models

Labeled data for NLP

The origin of natural language processing (NLP) dates back to the 70s, but thanks to developments in deep learning and computational resources, we now have the capacity to train NLP models on billions of pieces of supervised data to deliver richer features with increased levels of accuracy. Labeled data is the new oil, and enterprises need massive volumes of labeled data in order to stand up successful NLP applications.

While unsupervised learning has come a long way, supervised machine learning (ML) is critical for performing complex cognitive functions like deductive reasoning and inferences. The types of data a sophisticated NLP model needs to learn requires human-like frames of references that have not yet been recorded. And when the recorded data is not available, you need to consider building custom-labeled datasets.

Why NLP needs humans

NLP interactions require computers to understand and process the human language—so it should come as no surprise that it relies on human input. By leveraging the natural neural networks of the human brain to develop training data for NLP models, enterprises are able to realize its full potential.

Increasing confidence in your AI algorithms

Alegion works with enterprise brands to accelerate model training and validation through the delivery of high accuracy, high volume training data. Through our data labeling platform and a global network of specialists, we automate the human intelligence integration your models need to achieve your ML objectives. The result is high quality, human-augmented training datasets and compelling NLP-powered experiences.

1 — **2** — **3** — **4** — **5**

Solution Optimization
- Onboarding consultation with a dedicated CSM
- Task customization
- Workflow configuration

Task Management
- Specialized workforce curation, screening
- Global task distribution
- Workforce training, skills management

Data Labeling
- Text/audio/image labeling toolsets
- Text classification
- Semantic enrichment
- Entity recognition

Quality Control
- Judgment consensus
- Review, adjudicate, and escalate
- Confidence scoring
- Gold data

Model Validation
- Real-time routing for human review
- Model scoring and exception handling
- Ongoing solutions management

How Alegion accelerates your NLP projects

Unlike off-the-shelf alternatives, we take a solutions-first approach to enriching your training data by aligning ourselves to your business goals. Our platform integrates purpose-built task management tools and a network of qualified human specialists to label and quality control your data in a fraction of the time it would take to do so in-house. You reclaim the data labeling hours and compress your development cycle without having to engineer new tools, manage a workforce, or compromise on data quality.

- **Enterprise ML training best practices**
  de-risk projects and speeds time to value

- **Fully managed data labeling as a service**
  eliminates in-house labeling and project management

- **API systems integrations**
  support secure, real-time data exchange

- **ML-augmented quality management**
  ensures delivery of high-accuracy training data

- **Task workflows and annotation tools**
  are custom-configured for your project needs

- **Global network of distributed workforces**
  mitigates model training data bias

Chatbots
Capture complex queries into fixed classes to design conversations

Customer Service
Identify customer objectives with pattern recognition

Sentiment Analysis
Overlay human decisions to extract subjective information

Virtual Assistants
Help virtual assistants deliver personalized interactions

Mobile Apps
Keep NLP-driven features up-to-date with continuous model validation

Content Recommendation
Improve algorithms based on prior decisions

Information Extraction
Extract structured information from unstructured data

Intent Recognition
Create associational relationships to identify intent

Semantic Enrichment
Infer meaning and improve content utility through text augmentation

The smarter way to advance your NLP initiatives

Talk to us at **solutions@alegion.com** to learn how you can accelerate your model development and access high-quality training data in partnership with a Customer Success team who is as equally vested in the quality of your data output as you are.

Established in 2012, Alegion is an Austin-based AI acceleration company with a platform that integrates human intelligence into the creation of large-scale, high-quality **training data, model validation,** and **exception handling**. Alegion's platform serves Fortune 500 companies in **retail, finance, automotive, technology, government,** and **healthcare** industries to achieve **99%+ data accuracy** and a **Net Promoter Score® of 95.**

Learn more at alegion.com