Ariana Patel — NLP Engineer

Years of Experience: 13
PhD in Computer Science

ARIANA PATEL
NLP ENGINEER
San Francisco, CA - (555) 123-4567 - ariana.patel@example.com - LinkedIn: linkedin.com/in/ariana-patel

SUMMARY

Seasoned NLP engineer with 13 years of experience designing and deploying state■of■the■art language models for enterprise and consumer applications. Specialized in transformer architectures, large■scale data pipelines, and cross■functional team leadership. Adept at translating research innovations into production■ready solutions that drive measurable business value.

SKILLS

Natural Language Processing, Deep Learning, Machine Learning, PyTorch, TensorFlow, Transformers, BERT, GPT, spaCy, NLTK, Python, SQL, Git, Docker, Kubernetes, Data Pipeline Design

EXPERIENCE

- Developed enterprise■grade sentiment analysis engines for global brands, achieving >95% classification accuracy.
- Implemented real■time inference microservices with sub■100 ms latency, scaling to 10M+ requests/day.
- Mentored junior developers, establishing reproducible research workflows and coding standards.

- Managed cross■departmental NLP initiatives across finance, health, and retail, delivering 7+ production models.
- Directed migration from legacy RNN architectures to transformer■based models, reducing inference costs by 40%.
- Collaborated with ML Ops to deploy scalable inference pipelines on Kubernetes, ensuring zero■downtime releases.

- Oversaw a team of 4 engineers in end■to■end development of a multilingual chatbot used by 1M+ customers.
- Implemented transformer■based zero■shot classification for customer support tickets, boosting resolution time by 30%.
- Pioneered data■augmentation techniques that increased F1 scores by 12% on multi■label datasets.

- Designed and maintained entity extraction pipelines for legal documents, extracting over 1M entities per month.
- Built RESTful APIs for model inference using Flask and Docker, achieving 95% uptime.
- Optimized tokenization and vectorization steps, reducing inference latency by 35% compared to baseline.

- Led a project extracting sentiment signals from 200M+ social media streams daily.
- Integrated PyTorch■based BERT fine■tuning into a real■time analytics platform, improving sentiment recall by 18%.
- Collaborated with product managers to define model KPIs and report results to senior stakeholders.

- Assisted in developing contextual embedding models for large■scale corpora, contributing to a top■tier ACL paper.
- Implemented data preprocessing pipelines using Python and Spark, scaling data throughput by 2×.
- Authored a paper on attention mechanisms, presented at ACL 2014, receiving positive reviews from peers.

EDUCATION

Doctor of Philosophy (PhD), Computer Science - Massachusetts Institute of Technology, Cambridge, MA 2014

--