

## Day 13: Sentiment Analysis for Urdu & Pashto

Sami Uddin Shinwari  
Subject Specialist – IT | NLP Researcher  
MS Data Science (FAST NU Peshawar)  
BS Computer Science (UET Peshawar)

January 15, 2026

# What is Sentiment Analysis?

- Detecting opinions or feelings in text.
- Determines polarity: Positive, Negative, or Neutral.
- Widely used for social media, reviews, and feedback.

# Challenges for Urdu & Pashto

- Limited labeled datasets for training.
- Complex morphology affects word sentiment.
- Spelling variations and dialects.
- Lack of comprehensive sentiment lexicons.
- Borrowed words from Arabic, Persian, and English.

# Sentiment Example (English)

**Sentence:**

*I love this book!*

**Polarity: Positive**

# Sentiment Example (Urdu)

**Sentence:**

مچے یہ کتاب بہت پسند ہے

**Polarity: Positive**

# Sentiment Example (Pashto)

**Sentence:**

زه دا کتاب خوبیوم

**Polarity: Positive**

# Tools and Approaches

- Rule-based lexicon methods (rare for Urdu/Pashto)
- Machine learning models: SVM, Naive Bayes
- Transformers: mBERT, XLM-R (limited low-resource support)
- Need for domain-specific sentiment lexicons

# Key Takeaways

- Sentiment analysis is challenging for Urdu & Pashto.
- Morphology, spelling variations, and lack of lexicons are major hurdles.
- Deep learning can help but labeled data is scarce.
- Community-driven resources are crucial for progress.

**#Day13 #SentimentAnalysis #UrduNLP #PashtoNLP  
#LowResourceLanguages**