Day 1: Introduction to NLP

Sami Uddin Shinwari
Subject Specialist - IT — NLP Researcher
MS in Data Science (FAST NU Peshawar)
BS in Computer Science (UET Peshawar)

December 14, 2025

What is NLP?

- **Natural Language Processing (NLP)** is a branch of AI that allows machines to read, understand, and generate human languages.
- Powers applications like Google Translate, chatbots, sentiment analysis, and text classification.

Why NLP Matters for Pashto Urdu

- Most NLP research focuses on English; Pashto Urdu are **low-resource languages**.
- Challenges:
  - Complex scripts
  - Rich morphology
  - Dialectal variations
- Opportunities: Education, media, government, business applications.

Fun Fact

- Standard NLP tools for English often fail on Pashto Urdu.
- Tokenization, sentiment analysis, and text classification need **special handling**.

Stay Tuned!

- Follow this 30-day journey on NLP for Pashto, Urdu, and other languages!
- **NLP PashtoNLP UrduNLP AI LowResourceLanguages**