Our solution combines fast rule-based techniques with lightweight NLP models to turn raw chat transcripts into clear, actionable insights. First, the text is aggressively pre-processed and tokenised with SpaCy, from this we collect several complementary evidence streams—named entities, noun-phrases, adjective–noun and verb–noun compounds, plus problem-specific patterns such as “wrist pain”. A custom frequency–deduplication algorithm then ranks these candidates, boosting ergonomics- and pain-related terms, to output the “primary topics” in well under 100 ms.

Next, a weighted pattern-matching engine scores commercial intent. Positive signals (e.g. “I need…”, “can you recommend…”, urgency words) raise the score, while negatives (“already have…”, “don’t need…”) lower it; sentiment and conversation structure (number of user turns, recurring nouns, dominant emotions) add further context, producing a nuanced 0-1 intent value. Finally, each topic is mapped to product categories through a five-layer matcher: direct keyword dictionaries, category keyword overlap, regex patterns, domain-specific boosts, and related-category propagation. The result is returned alongside sentiment, urgency and rich “why” explanations, all rendered in a Streamlit dashboard that processes typical conversations end-to-end in under 200 ms.