**Self-Attachment Theory (SAT) Teacher Integration**:

Addressing the challenges faced by users unfamiliar with the self-attachment theory (SAT), we introduced an innovative feature: the SAT Teacher. This tool equips users to pose queries pertaining to the SAT, enabling them to grasp its central tenets, the associated exercises, and its overarching benefits. By leveraging this, even individuals who have not undergone an 8-week psychological intervention can gain insight into SAT, making our chatbot more accessible. Consequently, this enhancement expanded our user demographic, accommodating individuals with no prior knowledge of SAT.

To ensure comprehensive knowledge dissemination, we embedded over 140 question-answer pairs into our model. To optimize accuracy, each primary question was supplemented with two analogous questions. This entire dataset was initially rendered from English to Persian via the Google Translate API. Given that these translations were directly presented to users, meticulous revisions were conducted to rectify potential discrepancies, affirming the integrity of the presented content.

**Embedding and Similarity Computation**:

Utilizing the Pars-BERT model, we generated embeddings for each question. These embeddings facilitate an accurate understanding of question semantics. When a user query is received, our system transforms it into an embedding, capturing its essence.

Cosine similarity is employed to measure the alignment between the user's query and our dataset questions. The cosine similarity score, , between two vectors A and B is given by:

Where is the dot product of the vectors, and are the magnitudes of vectors and , respectively. Upon evaluating the cosine similarity scores for the user's query against our dataset, we select the highest score as the predicted match and present the associated answer to the user.

**Validation and Augmentation**:

As shown in Figure 6, we innovatively devised a methodology to generate semantically similar questions to our primary dataset. We capitalized on commonly used starter and ender phrases, which reflect user tendencies in chatbot interactions. Furthermore, to foster diversity and simulate real-world user queries, central words were paraphrased using their respective synonyms or semantically equivalent terms.

By deploying this strategy, we generated approximately four analogous questions for each principal query, culminating in over 500 similar questions. Remarkably, our model correctly predicted over 80% of these, a testament to its efficiency. Given the volume of primary and augmented questions, this precision rate underscores the robustness of our chatbot.

