Here is a structured set of instructions you can use as a knowledge base for an LLM-based semantic analyzer to judge sentiment and provide a sentiment score:

Instructions for Sentiment Analysis and Scoring

Step 1: Text Preprocessing

Before performing sentiment analysis, preprocess the text to ensure accuracy:

- Cleaning: Remove irrelevant characters, punctuation, emojis, and URLs.
- Tokenization: Break the text into individual words or tokens.
- Stop-word Removal: Remove common words that do not add significant meaning (e.g., "the," "is," "at," "which").
- Lemmatization or Stemming: Reduce words to their root form (e.g., "running" → "run")[4][7][8].

Step 2: Sentiment Classification Approaches

Use one of the following methodologies to classify sentiment:
Lexicon-based Approach

- Identify sentiment-bearing words using a predefined lexicon.
- Assign each word a sentiment value:
  - Positive words: +1
  - Negative words: -1
  - Neutral words: 0 (ignored)
- Handle negations explicitly (e.g., "not good" → negative sentiment).
- Calculate total sentiment by summing individual word scores[1][2][4].

Machine Learning-based Approach

- Train a supervised model (e.g., Logistic Regression, Naive Bayes, Support Vector Machines, Neural Networks) on labeled data.
- Classify unseen text based on learned patterns.
- Output classification as positive, negative, or neutral with associated confidence scores[1][2][5].

Transformer-based Deep Learning (LLM) Approach

- Use pretrained transformer models (e.g., BERT) fine-tuned on labeled sentiment data.
- Consider context and semantic nuances for accurate classification.

- Predict sentiment categories (positive, negative, neutral) along with probability/confidence scores[3][5].

Step 3: Sentiment Categories and Scoring Scale

Classify the analyzed text into the following sentiment categories with corresponding numerical scores:

Sentiment Category	Numeric Score	Description
Very Positive	+2	Strongly positive emotional tone
Positive	+1	Generally positive emotional tone
Neutral	0	Neither positive nor negative tone
Negative	-1	Generally negative emotional tone
Very Negative	-2	Strongly negative emotional tone
Mixed	0	Contains both positive and negative tones

Assign these numeric values based on the model's output probabilities or lexicon-based scoring[6].

Step 4: Calculation Methods for Numeric Scores (Lexicon-based)

If using lexicon-based scoring, choose one of these methods:

Method A: Simple Word Count

$$\text{Sentiment Score} = (\text{Positive Words Count}) - (\text{Negative Words Count})$$

$$\text{Sentiment Score} = (\text{Positive Words Count}) - (\text{Negative Words Count})$$

Method B: Normalized by Sentence Length

$$\text{Sentiment Score} = \frac{(\text{Positive Words Count}) - (\text{Negative Words Count})}{\text{Total Words in Sentence}}$$

$$\text{Sentiment Score} = \frac{(\text{Positive Words Count}) - (\text{Negative Words Count})}{\text{Total Words in Sentence}}$$

Method C: Ratio Method

$$\text{Sentiment Score} = \frac{\text{Positive Words Count}}{\text{Negative Words Count} + 1}$$

$$\text{Sentiment Score} = \frac{\text{Positive Words Count}}{\text{Negative Words Count} + 1}$$

Choose method based on your use case complexity and data length[3].

Step 5: Aspect-Based Sentiment Analysis (Optional)

For more granular analysis:

- Identify specific aspects or features mentioned in the text (e.g., product attributes).
- Separately analyze sentiment toward each identified aspect.
- Assign individual sentiment scores per aspect[5].