Algorithm 1: Medicine Information Classification

Input: Image file path
Output: Predicted medicine information

1. Import necessary libraries and modules;
2. Define `remove_special_characters()` function to remove special characters;
3. Define `get_first_word()` function to retrieve the first word;
4. Set Tesseract OCR executable path;
5. Open and process the input image using PIL and pytesseract;
6. Clean the extracted text by removing special characters;
7. Extract the first word from the cleaned text;
8. Read dataset from CSV file and handle missing values;
9. Separate input features and target variables;
10. Initialize label encoders and encode target variables;
11. Convert encoded target variables to DataFrame;
12. Split data into training and testing sets;
13. Tokenize text sequences in the training set using TensorFlow's Tokenizer;
14. Convert tokenized sequences to padded sequences;
15. Define dictionary to store trained models;
16. Iterate over target variables and create CNN models;
17. Compile and train models on the training data;
18. Prepare input medicine name by tokenizing and padding;
19. Make predictions using trained models and prepared medicine name;
20. Decode predicted purposes using label encoders;
21. Initialize engine for text-to-speech synthesis;
22. Iterate over predicted values and generate synthesized speech;
23. Print predicted values;