A. Data Description

1. The Aspect_terms folder includes files containing standard aspect terms, and the same or approximately standard aspect words on ontology are as follows:
   - ant_aspect.xls: safety aspect.
   - doc_aspect.xls: engine aspect.
   - gib_aspect.xls: price aspect.
   - ngt_aspect.xls: exterior aspect.
   - not_aspect.xls: interior aspect.
   - tot_aspect.xls: overall aspect.
   - vah_aspect.xls: performance aspect.

   Figure 1 shows the structure of these files.

	A	B	C
1	index	aspect_item	similar_item
2	1	an toàn	abs
3	2	an toàn	an toàn
4	3	an toàn	bám đường

Figure 1. Structure of files containing aspect terms.

The structure of the files is as follows:

- *index* column: identification of pattern.
- *aspect_item* column: the standard aspect terms.
- *similar_item* column: the same or approximately standard aspect words.

2. The word2vec_aspect folder includes the word2vec files. These files use for aspect detection tasks.
   - ant_word_size_300.model: safety aspect.
   - doc_word_size_300.model: engine aspect.
   - gib_word_size_300.model: price aspect
   - ngt_word_size_300.model: exterior aspect
   - not_word_size_300.model: interior aspect.
   - tot_word_size_300.model: overall aspect.
   - vah_word_size_300.model: performance aspect.

3. The Single_Aspect_Standardized_Sentence folder includes the files containing data used in training and testing that are standardized. Each sentence has only one aspect.
   - The data files for the engine aspect in the two experiments are as follows:
   - Engine_atempt_1.xls: The training data is the part from row 2 to row 34142.
   - Engine_atempt_2.xls: The training data is the part from row 2 to row 34132.
   - The data files for the safety aspect in the two experiments are as follows:
   - Safety_atempt_1.xls: The training data is the part from row 2 to row 34165.
   - Safety_atempt_2.xls: The training data is the part from row 2 to row 34160

- The data files for the price aspect in the two experiments are as follows:
- Price_atempt_1.xls: The training data is the part from row 2 to row 34061.
- Price_atempt_2.xls: The training data is the part from row 2 to row 34061.
- The data files for the exterior aspect in the two experiments are as follows:
- Exterior_atempt_1.xls: The training data is the part from row 2 to row 34015.
- Exterior_atempt_2.xls: The training data is the part from row 2 to row 34015.
- The data files for the interior aspect in the two experiments are as follows:
- Interior_atempt_1.xls: The training data is the part from row 2 to row 34161.
- Interior_atempt_2.xls: The training data is the part from row 2 to row 34142.
- The data files for the overall aspect in the two experiments are as follows:
- Overall_atempt_1.xls: The training data is the part from row 2 to row 33701.
- Overall_atempt_2.xls: The training data is the part from row 2 to row 33701.
- The data files for the performance aspect in the two experiments are as follows:
- Performance_atemp_1.xls: The training data is the part from row 2 to row 34074.
- Performance_atemp_2.xls: The training data is the part from row 2 to row 34086.

Figure 2 shows the structure of these files.

	A	B	C
1	id	text	polarity
2	200000	cái_giá an_toàn	other
3	200001	1 anh xe_ôm chia_sẻ hài_hước	other
4	200002	1 chiếc minivan đủ đảm_bảo an_toàn tính_năng kiểm_soát xe hiện_đại	other

Figure 2. Structure of files containing training and testing data for aspects.

The structure of the files is as follows:

- *id* column: identification of pattern.
- *text* column: the content of sentences.
- *polarity* column: sentiment labels.

4. The Single_Aspect_Raw_Sentence folder includes the files containing raw data used in training and testing. Each sentence has only one aspect.
   - The data files for the engine aspect in the two experiments are as follows:
   - Engine_atempt_1_raw.xls: The training data is the part from row 2 to row 8593.
   - Engine_atempt_2_raw.xls: The training data is the part from row 2 to row 8581.
   - The data files for the safety aspect in the two experiments are as follows:
   - Safety_atempt_1_raw.xls: The training data is the part from row 2 to row 8613
   - Safety_atempt_2_raw.xls: The training data is the part from row 2 to row 8610.
   - The data files for the price aspect in the two experiments are as follows:
   - Price_atempt_1_raw.xls: The training data is the part from row 2 to row 8413.
   - Price_atempt_2_raw.xls: The training data is the part from row 2 to row 8511.
   - The data files for the exterior aspect in the two experiments are as follows:
   - Exterior_atempt_1_raw.xls: The training data is the part from row 2 to row 8465.
   - Exterior_atempt_2_raw.xls: The training data is the part from row 2 to row 8465.
   - The data files for the interior aspect in the two experiments are as follows:

- Interior_atempt_1_raw.xls: The training data is the part from row 2 to row 8611.
- Interior_atempt_2_raw.xls: The training data is the part from row 2 to row 8613.

• The data files for the overall aspect in the two experiments are as follows:

- Overal_atempt_1_raw.xls: The training data is the part from row 2 to row 8152.
- Overal_atempt_2_raw.xls: The training data is the part from row 2 to row 8151.

• The data files for the overall aspect in the two experiments are as follows

- Performance_atemp_1_raw.xls: The training data is the part from row 2 to row 8593.
- Performance_atemp_2_raw.xls: The training data is the part from row 2 to row 8608.

These files have the same structure as shown in Figure 2.

5. The Data_for_sentiment_analysis folder contains the sentiment_data_train_3C_doc.xls file for the sentiment analysis task. The training data is the part from row 2 to row 37269. Figure 3 shows the structure of this file.

	A	B	C
1	id	text	polarity
2	1000	1 chiếc minivan đủ sang_trọng đảm_bảo an_toàn tiện_dụng	pos
3	1001	1 chiếc xe được bán chính hãng sinh nó đã vậy không gọi xe độ cả	neu
4	1002	$11m để làm chuyện khác hơn	neu
5	1003	1 chiếc xe hoàn_hảo	pos
6	1004	1 dàn lạnh người ngồi trước cóng người ngồi sau cuối sẽ mát	neg
7	1005	1 kiệt_tác của nhân_loại	pos

Figure 3. The structure of the sentiment analysis data file.

The structure of the files is as follows:

- *id* column: identification of pattern.
- *text* column: the content of opinions.
- *polarity*: sentiment label.

6. The Data_two_and_more_aspect_sentences_test folder contains the two_and_more_aspect_sentences_test.xls file for test multi-aspect in the aspect detection task. Figure 4 shows the structure of this file.

	A	B	C	D	E	F	G	H	I
1	id	sentences	safety	engine	price	exterior	interior	performance	overall
2	1001	S400 lên đến S600 nó khác nhau 1 trời 1 vực đấy từ động cơ nội đến ngoại thất như vậy có gọi là xe độ không	no	yes	no	yes	yes	no	no
3	1002	1 CHIẾC XE HOÀN HẢO NHƯNG HÀNG GHẾ SAU LÀ 1 THẢM HỌA CHO CÁI GIÁ 1O TY	no	no	yes	no	yes	no	yes
4	1003	1 5 tỷ mua con Sedona ngon lành hơn nhiều	no	no	yes	no	no	no	yes
5	1004	10 tỉ bỏ ra mua cái máy công nghệ của thế kỷ 19	no	yes	yes	no	no	no	no
6	1005	tội gì tầm giá này vay tiền đi mua Fortuner 27 mẫu hơi cũ nhưng như vậy lại đỡ nhàm chán phù hợp các bác trung niên thôi	no	no	yes	yes	no	no	no

Figure 4. The structure of the multi-aspect sentences data file.

The structure of the files is as follows:

- *id* column: identification of pattern.
- *sentences* column: the content of sentences.

- *safety, engine, price, exterior, interior, performance*, and *overall* column: determined "yes" or "no" aspect in sentence correspond.

7. The Opinion_for_summarize folder contains the opinion_test.xls file for test multi-aspect in the aspect detection task. Figure 5 shows the structure of this file.

	A	B	C	D	E	F	G	H	I
1	id	opinion	safety	engine	price	exterior	interior	performance	overall
2	10001	Thùng tôn di động xem ra không thay đổi nhiều, mà nhìn vẫn quá xấu.				neg			
3	10002	1 Bác chạy trên cao tốc 150 km / h thì chả có ý nghĩa gì, kia Morning hay bất kể con xe nào cũng làm được. Vấn đề là từ lúc depart đến khi đạt 150 km / h hết bao nhiêu thời gian. 2 Bác nói vượt xe khác chỉ nhíc cái là vượt, vấn đề là xe khác kia chạy tốc độ bao nhiêu? 3 Cùng là máy xăng 27 (Em tin chắc là cùng công nghệ chứ TOY nó chả đặt RR sản xuất động cơ cho con Prado) tại sao em thấy con Fortuner chạy ì thế? Chả có tỷ nào mạnh mẽ như bác nói về con Prado		pos					neu
4	10005	1 CHIẾC XE HOÀN HẢO NHƯNG HÀNG GHẾ SAU LÀ 1 THẢM HỌA CHO CÁI GIÁ 1O TY			neu	pos	neg		
5	10006	1 kiệt tác của nhân loại.							pos
6	10007	1 phiếu cho Honda							pos
7	10009	1,5 tỷ mua con Sedona ngon lành hơn nhiều.							neg
8	10010	10 tỉ bỏ ra mua cái máy công nghệ của thế kỷ 19 . máy 57 v8 mà có 368 mã lực thua máy của mec 30		neg					

Figure 5. The structure of the multi-aspect sentences data file.

The structure of the files is as follows:

- *id* column: identification of pattern.
- *opinion* column: the content of opinions.
- *safety, engine, price, exterior, interior, performance*, and *overall* column: indicate whether to review the respective aspect or not. The aspect that is reviewed in the opinion will be sentiment classifying with polarities as positive (pos), negative (neg), or neutral (neu).

B. Source code description

The Source_code folder contains source code files in Python that implement the proposed summary model. The main function of each file specification is as follows:

- data_url.py: declare global variables.
- helpers.py: the data processing functions.
- m_BiLSTM_multi_gpu.py: the BiLSTM model
- m_CNN.py: the CNN model
- m_CNN_BiLSTM.py: the CNN_BiLSTM model
- m_CNN_LSTM.py: the CNN_LSTM model
- m_data_aspect.py: the data processing for aspect functions.
- m_LSTM.py: the LSTM model.
- multi_opinion_summarize_final.py: multi-opinion summary system
- single_opinion_summarize_final.py: single-opinion summary system.
- CNN_sentiment_word_3c_relu.json: The configuration of the CNN model.
- CNN_sentiment_word_3c-039-0.3146-0.9361.h5: The feature of the corpus that the CNN model has trained.