Chapter 6

Aspect term category-wise sentiment aggregation framework for Hindi document

ABSA's two primary sub-tasks are aspect category detection(ACD) and aspect sentiment classification(ASC). The aspect category detection aims to identify the category of aspect present in review, aspect predicts the sentiment for the aspect. ABSA is a complex and thorough exercise for determining the sentiment expressed about certain textual parts. The aspect category is not explicitly mentioned in a review. A key component of opinion mining, a discipline concerned with drawing insightful conclusions from textual data, is ACD. There is a multi-label classification behind the aspect category sub-task. Review sentences are categorized into pre-established aspect groups. For example, in a restaurant evaluation, aspect categories like "FOOD", "SERVICE", "AMBIENCE" and so forth might be included in following review.

The pizza is horrible and the service is bad.

Once aspect category have been detected, the next task, i.e. ASC, predicts a sentiment class(e.g. *positive*, *negative*, *neutral* or *conflict*) for given review. $FOOD$ and $SERVICE$ are aspect categories. The costumer has expressed *negative* sentiment in this review. Two aspect categories $FOOD$ and $SERVICE$ are present implicitly in review, so this problem lies under multi-label classification. Negative sentiment is expressed in review.

Tasks involving multi-label classification encompass diverse methodologies, including issue transformation and algorithm adaptation. Problem transformation is utilized to split the multi-label classification problem into multiple discrete single-label problems. In the meantime, specific algorithms are modified in the problem adaption approach to handle multi-label classification jobs efficiently. Simultaneously a multi-class classification algorithm is needed to predict sentiment expressed.

In this chapter we develop a novel multi-task learning framework. Our objective is to simultaneously identify aspect categories and predict aspect category phrases in a single model.

6.1 Methodology

The suggested method for multi-task learning for aspect sentiment categorization and aspect category detection is discussed in this section.We consider the aspect term category wise sentiment aggregation as a multi-task learning. The aspect category detection is a multi-label classification(MLC) and second, the aspect sentiment prediction is a multi-class classification(MCC). The two assignment are completed in order. We examined binary relevance, classifier chain and label power-set approaches for aspect category detection. Label power-set approach provides significantly better results[46]. Support Vector Machine(SVM), Random Forest(RF), Logistic Regression(LR), XG-Boost(XGB), k-Nearest Neighbour(KNN) and Naive Bayes(NB) classifiers are examined for MLC and MCC Models with TF-IDF and Count Vectorization method. Evaluation results are presented in Section 6.2.1. XG-Boost classifier with count Vectorization produces the most pertinent results for MLC and MCC models. The complete work flow is presented in fig 6-1. Moreover, we have outlined the processes that comprise the

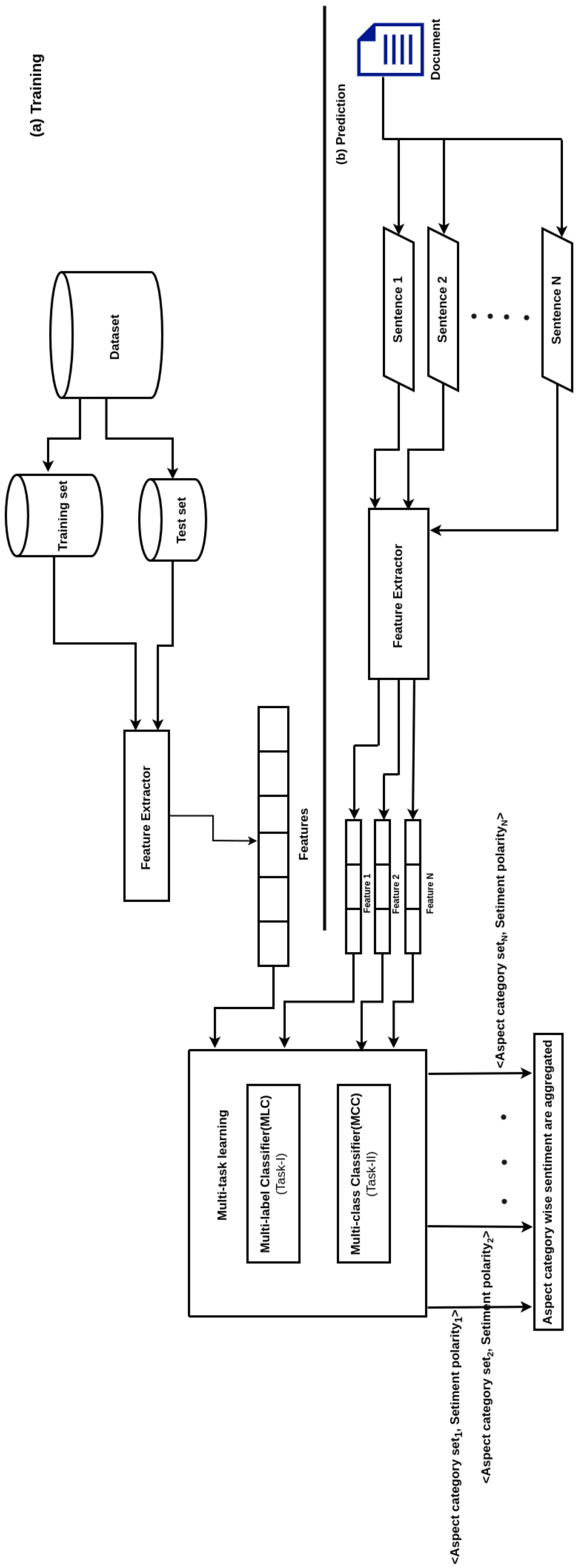

Figure 6-1: Multitask learning workflow

complete procedure as follows:

1. Apply following steps to dataset IITP-II to make it appropriate for multi-label assignment and multi-class assignment:

   (a) Initialize an empty list to store Hindi sentence data and aspect category labels.

   (b) Store each sentence in one row.

   (c) Store corresponding polarity in one row.

   (d) Initialize all aspect categories (AC_1, AC_2, AC_3..... AC_m) by 0 for each sentence.

2. For each sentence:

   (a) Identify and collect aspect categories.

   (b) Update the set of aspect categories with 1 for the present aspect categories.

3. Split dataset IITP-II in training set T_r and test set T_s.

   #Classifier I(MLC)

   (a) Drop polarity column from training set T_r and test set T_s.

   (b) Perform feature extraction on column *Text* of training set T_r and test set T_s using TF-IDF/Count Vectorization.

   (c) Apply multi-label classifier on T_r and T_s.

   (d) Apply 10-fold cross validation.

   #Classifier II(MCC)

   (a) Drop polarity column from training set T_r and test set T_s.

   (b) Perform feature extraction on column *Text* of training set T_r and test set T_s using TF-IDF/Count Vectorization.

   (c) Apply multi-class classifier on T_r and T_s.

(d) Apply 10-fold cross validation.

4. Aspect category and sentiment vector generation.

   Input : A document(D)

   (a) Apply sentence-ending punctuation to split the document into sentences.
   $D = S_1, S_2, S_3, S_4, S_5, ...S_n$
   n = number of sentences.

   (b) For each sentence S_i where $1<=i<=n$, perform preprocessing.
   $S_i^{'}$:= pre-process(S_i).

   (c) Perform feature selection for each pre-processed sentence($S_i^{'}$), Let $FS_1^{'}$, $FS_2^{'}$, ..., $FS_n^{'}$ be the feature vectors for the n sentences.

   (d) Apply the trained MLC to each $FS_i^{'}$.

   (e) Record the predicted aspect categories($AC_1^{'}$, $AC_2^{'}$, $AC_3^{'}$, ... $AC_m^{'}$) for each $S_i^{'}$.

   (f) Apply the trained MCC to each $S_i^{'}$.

   (g) Record the predicted polarity (PP) for $S_i^{'}$.
   If PP is *pos*, set it to numerical value 1;
   if PP is *neg*, set it to numerical value -1;
   if PP is *neu*, set it to numerical value 0;

   (h) Append $AC_1^{'}$, $AC_2^{'}$, $AC_3^{'}$, ... $AC_m^{'}$ with PP.

   Output : Vectors V_1, V_2, V_3, ...V_n of length (m+1) .

   ACD lies under multi-label classification paradigm. ASC lies under multi-class classification paradigm.

5. Aspect Category and Sentiment polarity of Document

   Input : Vectors V_1, V_2, V_3.......V_n; number of aspect categories : m.

(a) Set each Final aspect Category FAC_1, FAC_2, FAC_3 ..... FAC_m to numerical value 0.

(b) Set p=m+1

(c) From $i = 1$ to m:

For each vector V_j where $1 <= j <= n$.

if ($V_j[i]$ == 1):

$FAC_i = FAC_i + V_i[p]$

Output : FAC_1, FAC_2, FAC_3, ..., FAC_m.

Observations:

For an aspect category AC_i where $1 <= i <= m$

1. if $FAC_i > 1$ Document D is positive for aspect category i
2. if $FAC_i == 0$ Document D is neutral for aspect category i
3. if $FAC_i < 1$ Document D is negative for aspect category i

6.2 Results and Analysis

We test our model on the IITP-II dataset. Detailed statistics and labelled structure for IITP-II dataset is presented in chapter 2. Following steps shows an illustration for a given Hindi document. IITP-II dataset is utilized for training.

Step-1

As the first step, dataset IITP-II is transformed as shown in figure 6-2. This step is multi-label assignment step. The steps of multi-label assignment are discussed in Chapter 5.

ID	Polarity	Text	Electronics	Misc	Movies	Travels
sma_299	pos	गियर 2 स्पोर्ट्स मे 1.6 इंच का सुपर एमोलेड डिस्प्ले लगा है जो आपको उम्दा रंग व आउटडोर विजिबिलिटी देने मे सक्षम साबित होगा।	1	1	0	0
sma_300	pos	मेटालिक बॉडी से बनी यह घड़ी आपको एक टिकाऊ सुविधा देने का वादा करती है।	1	0	0	0

Figure 6-2: Dataset snippet

Here total number of aspect categories(m) is four i.e. Electronics(AC_1), Misc(AC_2), Movies(AC_3) and Travels(AC_4).

$Document(D)$:= इसमें मेटेरियल इन्स्पायर्ड डिजाइन है जो एंड्रॉयड यूजर्स को मॉडर्न लगेगा। यह बच्चों और लेडीज के लिए शानदार कैमरा है। जबकि फुजीफिल्म मिनी 25 एक शानदार सेल्फी कैमरा है।

Step-2

Apply multi-task learning MLC and MCC to IITP-II.

Step-3

Aspect category and sentiment vector generation.

- S_1:= इसमें मेटेरियल इन्स्पायर्ड डिजाइन है जो एंड्रॉयड यूजर्स को मॉडर्न लगेगा।
- S_2:= यह बच्चों और लेडीज के लिए शानदार कैमरा है।
- S_3:= जबकि फुजीफिल्म मिनी 25 एक शानदार सेल्फी कैमरा है।

Total number of sentences(n):= 03

For each sentence S_i, pre-processing is performed. $S_1^{'}$, $S_2^{'}$ and $S_3^{'}$ are per-processed sentences.

- $S_1^{'}$ = मेटेरियल इन्स्पायर्ड डिजाइन एंड्रॉयड यूजर्स मॉडर्न लगेगा
- $S_2^{'}$ = बच्चों लेडीज शानदार कैमरा
- $S_3^{'}$ = फुजीफिल्म मिनी 25 शानदार सेल्फी कैमरा

Apply TF-IDF/Count Vectorizer to each sentence, $FS_1^{'}$, $FS_2^{'}$, $FS_3^{'}$ are feature vectors for pre-processed sentences S_1', S_2' and S_3'.

Apply MLC to FS_1', FS_2', FS_3'.

AC_1'=[1 0 0 0]

AC_2'=[0 1 0 0]

AC_3'=[1 0 0 0]

Apply MCC to FS_1', FS_2', FS_3'.

$PP_1 = neu : 0$

$PP_2 = pos : 1$

$PP_3 = pos : 1$

Append [1 0 0 0] and 0;

V_1 = [1 0 0 0 0]

Append [0 1 0 0] and 1;

V_2 = [0 1 0 0 1]

Append [1 0 0 0] and 1;

V_3 = [1 0 0 0 1]

Following are interpretations of V_1, V_2 and V_3.

V_1 = [1 0 0 0 0]

S_1 = इसमें मेटेरियल इन्स्पायर्ड डिजाइन है जो एंड्रॉयड यूजर्स को मॉडर्न लगेगा।

Electronics(1), Misc(0), Movies(0), Travels(0), ***NEU***

Aspect category for Sentence S_1 is Electronics and aspect polarity is ***NEUTRAL***.

V_2 = [0 1 0 0 1]

S_2 = यह बच्चों और लेडीज के लिए शानदार कैमरा है।

Electronics(0), Misc(1), Movies(0), Travels(0), ***POS***

Aspect category for Sentence S_2 is Misc and aspect polarity is ***POSITIVE***.

Append [1 0 0 0] and 1;

V_3 = [1 0 0 0 1]

S_3 = जबकि फुजीफिल्म मिनी 25 एक शानदार सेल्फी कैमरा है।

Electronics(1), Misc(0), Movies(0), Travels(0), ***POS***

Aspect category for Sentence S_2 is Electronics and aspect polarity is ***POSITIVE***.

Step-4

Aspect Category and Sentiment polarity for Document(D)

$V_1 = [1\ 0\ 0\ 0\ 0]$

$V_2 = [0\ 1\ 0\ 0\ 1]$

$V_3 = [1\ 0\ 0\ 0\ 1]$

(a) $FAC_1 = 0$, $FAC_2 = 0$, $FAC_3 = 0$,$FAC_4 = 0$

(b) $p = 4 + 1 = 5$

(c) $m = 4$

From $i = 1$ to 4:

For each vector V_j where $i <= j <= n$

if ($V_j[i]$ `==` 1):

$FAC_i = FAC_i + V_i[p]$

For $i = 1$, $j = 1$

$V_1[1] == 1$; $FAC_1 = FAC_1 + V_1[5] = 0 + 0 = 0$,

For $i = 1$, $j = 2$

$V_2[1] == 0$; no change in FAC_1

For $i = 1$, $j = 3$

$V_1[3] == 1$; $FAC_1 = FAC_1 + V_3[5] = 0 + 1 = 1$,

$FAC_1 = \mathbf{1}$ For $i = 2$, $j = 1$

$V_1[2] == 0$; no change in FAC_2

For $i = 2$, $j = 2$

$V_2[2] == 1$; $FAC_2 = FAC_2 + V_2[5] = 0 + 1 = 1$,

For $i = 2$, $j = 3$

$V_3[2] == 0$; no change in FAC_2

$FAC_2 = \mathbf{1}$ For $i = 3$, $j = 1$

$V_1[3] == 0$; no change in FAC_3

For $i = 3$, $j = 2$

$V_2[3] == 0$; no change in FAC_3

For $i = 3$, $j = 3$

$V_3[3] == 0$; no change in FAC_3

FAC_3 $=$ **0** For $i = 4$, $j = 1$

$V_1[4] == 0$; no change in FAC_4

For $i = 4$, $j = 2$

$V_2[4] == 0$; no change in FAC_4

For $i = 4$, $j = 3$

$V_3[4] == 0$; no change in FAC_4

FAC_1=1, FAC_2=1, FAC_3=0, $FAC_4 = 0$

Observations

Document(D) is positive for aspect category FAC_1 (Electronics),

Document(D) is positive for aspect category FAC_2 (Misc),

Document(D) is neutral for aspect category FAC_3 (Movies),

Document(D) is neutral for aspect category FAC_4 (Travels).

6.2.1 Performance Analysis

There are two classifiers used; One classifier, MLC is used for multi-label classification of aspect category. One classifier, MCC is employed for multi-class classification of aspect sentiment polarity. Performance analysis of MLC and MCC is presented in below.

6.2.1.1 MLC

The classifier derives the features from the input Hindi data set and applies the label power-set transformation method. It assigns the predefined aspect categories to the aspect terms in the given review. Among the machine learning classifiers Gaussian Naive Bayes(GNB), Multinomial Naive Bayes(MNB), Random forest(RF), Logistic regression(LR),k-nearest neighbours(KNN), Support Vector Machine(SVM) and XGBoost(XGB) are examined. The classifiers mentioned above are used with label power-set transformation technique. TF-IDF and count vectorization methods are used for feature selection. Performance of MLC is evaluated by utilizing performance metrics i.e. precision, recall and F1-score.

Table 6.1: Evaluation metrics for MLC

Model	**Feature Extraction Method**	**Metrics**	**SVM**	**RF**	**LR**	**XG**	**KNN**	**NB**
MLC	TF-IDF	Precision	0.69	0.64	0.86	0.63	0.67	0.67
		Recall	0.73	0.67	0.50	0.67	0.69	0.68
		F1-score	0.71	0.65	0.63	0.65	0.68	0.67
		Accuracy	0.66	0.60	0.51	0.60	0.63	0.60
	Count Vectorization	Precision	0.70	0.71	0.73	**0.73**	0.58	0.67
		Recall	0.71	0.75	0.73	**0.76**	0.61	0.68
		F1-score	0.71	0.73	0.73	**0.74**	0.59	0.68
		Accuracy	0.65	0.68	0.68	**0.69**	0.54	0.61

Table 6.1 show the performance of label power-set classifiers when applied using the TF-IDF and count vectorization methods. The XG-Boost classifier achieved the best results among all the label power-set classifiers. It showed highest precision 0.63, 0.67 recall, 0.65 F1-score and 0.60 accuracy with TF-IDF feature extraction and highest precision 0.73, 0.76 recall, 0.74 F1-score and 0.69 accuracy with Count Vectorization feature extraction. Each of the classifiers exhibits good precision, recall and F1-score.

The Logistic Regression (LR) model achieves the highest precision with TF-IDF feature extraction, indicating its strong capability in correctly identifying positive instances. However, its lower recall suggests that it misses a significant number of relevant instances, highlighting a potential trade-off between precision and recall. In contrast, the Count Vectorization method generally yields better performance for the MLC model, particularly with XGBoost (XG), which attains the highest F1-score. This indicates a more balanced ability to identify both positive and negative instances compared to TF-IDF.

6.2.1.2 MCC

MCC predicts the sentiment polarity of a given review. The machine learning techniques LR, DT, SVM, RF, KNN, GNB, MNB and XGBoost are employed. Feature selection is done using count-vectorization and TF-IDF algorithms. Using performance indicators, such as precision, recall, and F1-score, the effectiveness of MCC is assessed.

Table 6.2: Evaluation metrics for MCC

Model	**Feature Extraction Method**	**Metrics**	**SVM**	**RF**	**LR**	**XG**	**KNN**	**NB**
MCC	TF-IDF	Precision	0.63	0.76	0.71	0.62	0.53	0.52
		Recall	0.54	0.49	0.51	0.52	0.48	0.51
		F1-score	0.55	0.47	0.49	0.52	0.48	0.51
		Accuracy	0.62	0.59	0.63	0.61	0.51	0.56
	Count Vectorization	Precision	0.57	0.65	0.62	**0.65**	0.46	0.53
		Recall	0.56	0.55	0.58	**0.58**	0.43	0.53
		F1-score	0.57	0.59	0.59	**0.59**	0.41	0.53
		Accuracy	0.61	0.65	0.64	**0.66**	0.49	0.57

The evaluation of the MCC model's performance using different feature extraction methods reveals critical insights. With TF-IDF, the Random Forest (RF) model demonstrates the highest precision, indicating its effectiveness in correctly identifying positive instances. However, its recall is lower, suggesting that it misses many relevant instances, which may be concerning in applications where capturing all positive cases is crucial.

Table 6.3: Comparison of accuracy for multitasking model with English

Author(s)	**Language**	**Accuracy**
Cai et al. [21]	English	85.3%
Li et al. [69]	English	82.5%
Our Method	Hindi	69.7%

In contrast, when using Count Vectorization, XGBoost (XG) achieves the highest precision and recall, resulting in the best F1-score. This highlights balanced performance of XGBoost classifier, making it suitable for scenarios requiring both accuracy and comprehensive identification of positive cases. Additionally, Logistic Regression (LR) shows relatively stable performance with both methods, but its metrics indicate that improvements could be made to enhance its overall effectiveness.

6.3 Summary

Our approach introduces a novel multi-task learning framework that integrates aspect category detection and aspect sentiment classification into a single model. The process involves transforming the dataset for multi-label and multi-class assignments, training classifiers for each task, and generating vectors representing aspect categories and sentiment polarities for document aggregation.

For evaluation, the IITP-II dataset was used, demonstrating that our framework accurately categorizes and assigns sentiments to aspects within Hindi documents. Multi-label and multi-class classifiers show high precision, recall, and F1-scores, with XGBoost classifier and Logistic Regression performing particularly well.