ABSTRACT

With the gradual expansion of the work order data of the electric customer service system, the reliance on manual experience has led to low processing timeliness, unable to effectively discover the true demands of customers, and thus unable to provide customers with high-quality solutions. Most of the current methods use deep learning techniques such as Principal Component Analysis and Neural Networks to analyze the semantics of work orders, but it is difficult to fully capture the semantic information hidden in the work order title and work order description, which leads to a decrease in the performance of solution recommendation. Therefore, this paper proposes a Customer Preference-Aware Customer Service Solutions Recommendation (CPACS) model. In order to enhance the representation of work order data, the model uses customer preference information to “query” work order title sequence and work order description sequence separately, generates the temporal dependency representation of these two sequences, and then uses 1-D CNN and Transformer to capture the local and global temporal dependency information of these two sequences. Then, a new information fusion method, Additive Conv-Transformer Skip (ACT-Skip), is proposed to fuse the local and global dependency information in the work order data to improve the solution recommendation performance. The final experiments show that the CPACS model can perform representation learning on work order data more effectively than the baseline model, thus realizing superior performance in the customer solution recommendation task.

CCS CONCEPTS

• Computing methodologies → Artificial intelligence; • Information systems → Information systems applications; Data mining.

KEYWORDS

work order data, customer preference, solutions recommendation

1 INTRODUCTION

The business of power grid enterprises is complex. There is a large amount of semi-structured and unstructured text data in the production and operation of the power grid enterprise, involving a variety of professional fields such as electronics, chemistry, mechanics, information, etc. For example, the maintenance report in electric power contains professional equipment related to machinery, chemistry, physics, electronics and other professional field knowledge. This type of text data belongs to low-density value data, which has the characteristics of large data volume, complex structure, and lack of standardization. It is one of the difficult areas for data analysis and mining at present.

The work order data in the electric power customer service system belongs to the above-mentioned typical data. They mainly adopt the form of spoken description and record a large number of electric power business characteristics, but at the same time, the text also contains many electric power technical terms. The format of this type of text data is not uniform and the content varies greatly. At present, the work order content is processed and classified mainly through the judgment of the staff. Due to manual experience, the processing timeliness is low, the classification rules are inconsistent, and the true demands of customers cannot be effectively discovered. Therefore, there is an urgent need for an effective text data mining method that can automatically analyze

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

ICBDT 2021, September 24–26, 2021, Zibo, China
© 2021 Association for Computing Machinery. ACM ISBN 978-1-4503-8509-1/21/09... $15.00
https://doi.org/10.1145/3490322.3490329

Shuangshuang Guo∗
State Grid Shandong Electric Power Company Information and Communication Company, Jinan, China
guoshuangshuang@sd.sgcc.com.cn

Mingli Yin
State Grid Shandong Electric Power Company, Jinan, China
yimingli@sd.sgcc.com.cn

Wenbin Zhang
State Grid Shandong Electric Power Company Information and Communication Company, Jinan, China
wbzhang@sd.sgcc.com.cn

Ming Xu
State Grid Shandong Electric Power Company, Jinan, China
xuming@sd.sgcc.com.cn

Yujiang Wang
State Grid Shandong Electric Power Company Information and Communication Company, Jinan, China
wangyujiang@sd.sgcc.com.cn

Yuliang Shi
School of Software, Shandong University, Jinan, China; Dareway Software Co., Ltd, Jinan, China
shiyuliang@sdu.edu.cn

∗Shuangshuang Guo is the corresponding author.
the work order data and further explore the potential demands of customers, which is of great significance to the improvement of the level of marketing customer service.

Artificial intelligence and text mining technology have gradually been applied to various scenarios of electric power. Ma et al. [1] used the KNN algorithm to classify and calculate defective texts to construct a comprehensive evaluation model of circuit breaker status. Liu et al. [2] introduced the LDA topic generation model, and combined the content of the work order to construct a work order text mining model to filter, judge and classify the work order. Zou et al. [3] established the customer service work order automatic classification model to realize the rapid classification of work orders and dig out the important information in them, so as to provide a basis for analyzing the demands of users. Based on the above research work, it can be found that most of the existing methods use Principal Component Analysis (PCA) [4], topic model [5] or calculation of TF-IDF [6] and other methods for semantic analysis of topic or word-level work orders. These methods usually stay in the representation learning of the shallow features of the text, but lack fine-grained and deep semantic understanding. This shallow representation method is difficult to fully capture the hidden semantic information contained in the work order description, which leads to the degradation of solution recommendation performance. Therefore, we need to have an accurate analysis of the work order title and work order description in order to more completely express the semantic information contained in them and achieve accurate solution recommendation.

In response to the above problems, this paper proposes a Customer Preference-Aware Customer Service Solutions Recommendation (CPACS) model. The main contributions are as follows:

- A dot product attention mechanism is used to enhance the presentation of the title information and description information in the work order, and use dynamic preference information to "query" and reconstruct the title information and description information in the work order data.
- We use 1D CNN and Transformer to capture the local temporal dependence and global temporal dependence information in the work order data.
- We propose an Additive Conv-Transformer Skip (ACT-Skip) information fusion method, and use ACT-Skip to fuse local and global timing dependencies in work order data, thereby improving the recommendation effect of the model.

2 RELATED WORK

With the gradual expansion of the data accumulated in the power system, research on sentiment analysis and work order classification based on work order data is currently mainly carried out.

2.1 Sentiment Analysis

The main goal of sentiment analysis technology is to identify the emotional information expressed by users based on text data, and divide the text data into positive and negative types. Liu et al. [13] proposed a multi-sentiment analysis framework that integrates active learning and multiple supervised learning for joint voting. Xi et al. [7] proposed a two-stage domain sentiment dictionary construction algorithm based on the point mutual information and context constraints between sentiment words, which improved the ability to recognize the sentiment tendency of sentiment words. Zhong et al. [8] studied the text classification algorithm based on matrix projection (MP) and normalized vector (NLV) to realize sentiment analysis of product evaluation. Liu et al. [9] took word-level vector and word-level vector as original features, and then used convolutional neural network to extract text features and conduct sentiment orientation analysis. The results showed that word level vectors can achieve higher accuracy.

2.2 Work Order Classification

Lin et al. [10] used the principal component analysis method to analyze the relationship between the specific work order content and the time dimension, and realized the optimization of the customer service work order classification model. Wang et al. [11] proposed an LDA-based classification model for hot business work orders, which realized classification and screening of hot business work orders by learning the topic of the work order text. Zou et al. [3] first preprocessed the text in the electric power work order, and obtained the text representation by calculating the TF-IDF value, and finally used the decision tree model to realize the work order classification. With the rise of neural network language models, Xie et al. [6] used the Word2vec model to learn the word vector representation in the work order text, and then obtained the semantic representation of the text for work order classification [12].

In response to the above research, this paper constructs a customer preference-aware solution recommendation model based on the work order title and work order description to achieve adaptive matching of common problems and solutions.

3 METHOD

Figure 1 shows the overall process of CPACS model proposed in this paper. The model first uses the customer’s dynamic preference information to enhance and enrich the title information and description information of work order data, and then uses 1D-CNN and Transformer to capture the local and global timing dependence information in the work order data. In addition, an Additive Conv-Transformer Skip (ACT-Skip) method is used to fully integrate local and global dependent information. Finally, based on the fusion of the title information representation and the description information representation, we use the fully connected operation to recommend customer solutions.

3.1 Text Data Preprocessing

In the work order data of the customer service system, the question sentence is composed of several words or phrases. To obtain the key information in the customer question sentence, the question sentence needs to be segmented [14]. For example, the question "how to deal with network failure?", after word segmentation "how to / r deal with / v network / n failure / n ? / x". Through the research and analysis of the related technology of word segmentation and the characteristics of question sentences in the customer service system of this paper, we decide to use the jieba word segmentation method [14] as the word segmentation tool to segment the work order data. At the same time, for domain-specific words, these
where \(d\) is the adjustment coefficient, usually taken as 0.85; \(V\) represents the number of work orders. The \(|\text{Out}(V_j)|\) refers to the set of numbers pointed to by the \(V_j\) node; \(|\text{Out}(V)|\) refers to the number of nodes in the set.

### 3.2 Keyword Feature Extraction

In the customer service system, extracting the feature words in the question is a very important step. The feature word extraction refers to the use of related technologies to obtain the core words in the question that can represent the question. In this paper, a feature word extraction method based on graph model (TextRank algorithm) is used to extract key feature words. The algorithm is a sorting algorithm based on graphs. Its basic idea is derived from Google’s PageRank algorithm. It treats the text as composed of several words and establishes the corresponding graph model, and then uses the voting mechanism to rank the important words in the text, so that the feature words can be extracted only by relying on the structural relationship of the text itself. This method is simple, effective and widely used. The weight iteration formula of this algorithm is shown in the formula:

\[
S(V_i) = (1 - d) + d \times \sum_{j \in \text{In}(V_i)} \frac{1}{|\text{Out}(V_j)|} S(V_j)
\]

Where \(d\) is the adjustment coefficient, usually taken as 0.85; \(\text{In}(V_i)\) represents the set of all nodes pointing to the \(V_i\) node, and \(\text{Out}(V_j)\) refers to the set of nodes pointed to by the \(V_j\) node; \(|\text{Out}(V)|\) refers to the number of nodes in the set.

### 3.3 Recommendation Model

In this paper, the data consists of three parts: work order title, work order description, and customer information. Based on the described methods in subsections 3.1 and 3.2, we perform word segmentation and keyword feature extraction for the three parts of data respectively. Then we use embedding vectors for embedding representation learning.

Define the work order title sequence \(L \in \mathbb{R}^{T \times M}\), the work order description sequence \(D \in \mathbb{R}^{T \times N}\), and the customer’s dynamic preference information sequence \(P \in \mathbb{R}^{T \times S}\), where \(M, N,\) and \(S\) respectively represent the number of words in the longest sentence in the title sequence, preference sequence and work order description sequence. \(T\) represents the number of work orders. The CPACS model first learns an embedded representation of the above information, as shown below:

\[
E_L = LW_L \\
E_P = PW_P \\
E_D = DW_D
\]

Among them, \(W_L \in \mathbb{R}^{M \times \text{embedding}}, W_P \in \mathbb{R}^{N \times \text{embedding}},\) and \(W_D \in \mathbb{R}^{M \times \text{embedding}}\) are the weight matrices.

We introduce the scaled dot product attention mechanism (Scaled Dot-Product Attention) to learn the relationship \(R_L\) between customer preferences and work order title sequence, and the relationship \(R_D\) between customer preferences and work order description sequence, so as to use customer preference information to generate the representations of work order title information and description information, \(\text{Norm()}\) represents the normalization operation of the layer:

\[
R_L = \text{Norm} \left( \text{softmax} \left( \frac{E_P E_L^T}{\sqrt{d_{\text{embedding}}}} \right) E_L \right)
\]

\[
R_D = \text{Norm} \left( \text{softmax} \left( \frac{E_P E_D^T}{\sqrt{d_{\text{embedding}}}} \right) E_D \right)
\]

Next, this paper uses a one-dimensional convolution operation to mine the local temporal dependency information in the work order title sequence and the work order description sequence, respectively, and generate local contexts \(C_L\) and \(C_D\). In this paper, we set the size of the convolution kernel to \(2, d_{\text{embedding}}\):

\[
C_L = \text{Conv}_1(D(R_L))
\]

\[
C_D = \text{Conv}_1(D(R_D))
\]

In addition, the Transformer encoder module is used to generate a contextual information representation of the global work order title sequence and the work order description sequence. The structure of Transformer is shown in Figure 2.

\[
Q_L' = K_L' = V_L' = E_L
\]

\[
\text{Head}_{Li} = \text{softmax} \left( \frac{(Q_L' W_{E_L}^Q) (V_L' W_{E_L}^K)^T}{\sqrt{d_{\text{embedding}}}} \right) V_L' W_{E_L}^V
\]
Where $L = \text{work order data sequence}$ is used as input, and the final recommendation result is output through the multi-layer perceptron (MLP): $O_L = \text{Norm(FFN (Norm (MultiHead (Q_L', K_L', V_L') + E_L))}$

$+\text{Norm (MultiHead (Q_L', K_L', V_L') + E_L))}$

Among them, $W_L^O, W_L^K, W_L^V$ represent the weight parameter matrix, and $\text{FFN()}$ represents a layer of feedforward neural network.

$$Q_L' = K_D = V_D = E_D$$

$$Head_{Di} = \text{softmax} \left( \frac{(Q_D' W_D^O) (V_D' W_D^K)}{\sqrt{d_{\text{embedding}}}} \right) V_D' W_D^V$$

$$MultiHead (Q_L', K_L', V_L') = \text{Concat} (Head_{Di}, Head_{Di}, ..., Head_{Di}) W_L^O$$

$$O_D = \text{Norm(FFN (Norm (MultiHead (Q_D', K_D', V_D') + E_D))}$$

In order to fully combine the obtained global context information representation with the information representation of the local data, we design an information fusion method called Additive Conv-Transformer Skip (ACT-Skip). The specific process is shown in Figure 1. The calculation of the comprehensive representation vector Out of the work order data sequence is as follows:

$$Out = \text{Norm(Concat (ACT - Skip ([C_L, O_L]), ACT - Skip ([C_D, O_D])))}$$

$$ACT - \text{Skip} ([C_L, O_L]) = O_L + \sigma (C_L + O_L) + (1 - \sigma (C_L + O_D)) + C_L$$

$$ACT - \text{Skip} ([C_D, O_D]) = O_D + \sigma (C_D + O_D) + (1 - \sigma (C_D + O_D)) + C_D$$

Where $\sigma ()$ represents the sigmoid function.

Finally, the obtained comprehensive representation Out of the work order data sequence is used as input, and the final recommendation result is output through the multi-layer perceptron (MLP):

$$y = \text{Softmax(MLP(Out))}$$

### 4 EXPERIMENT

#### 4.1 Data

In order to verify the effectiveness of the model proposed in this article, we selected data obtained from an electric power customer service system in China, including 16259 work order data for 3638 customers from May 1, 2021 to May 31, 2021. This data contains data information such as the work order title, work order description, customer information, and corresponding solutions.

#### 4.2 Experiment Description

The experimental environment is shown in Table 1.

<table>
<thead>
<tr>
<th>Environment setting</th>
<th>Operating system</th>
<th>RAM</th>
<th>CPU</th>
<th>GPU</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Ubuntu</td>
<td>256GB</td>
<td>AMD 3970X</td>
<td>RTX 3090 x2</td>
</tr>
</tbody>
</table>

The experimental results in Table 2 show that the performance of the model proposed in this paper is better than other baseline models on all evaluation indicators. Because the model in this paper is based on customer information, it learns customer preference information from two aspects of work order title and work order description, thereby improving the recommendation performance of the model. The experimental results in Table 2 verify the effectiveness of the model proposed in this paper.

#### 4.3 Performance Comparison

In order to study the parameter sensitivity of the model and explore the recommendation performance of the model under different embedding dimensions ($d_{\text{embedding}}$), this paper conducts a parameter search experiment. The result is shown in Figure 3. It can be seen from the results in the figure 3 that when $d_{\text{embedding}} = 128$, the performance of the model reaches the maximum. Therefore, this paper sets the embedding dimension to 128.

#### 4.4 Ablation Experiment

Table 4 describes the comparison between the model variants. From Table 4, we can conclude that the performance of the model proposed in this paper is better than the variants of the CPACS model. At the same time, it also shows that every part of the CPACS model is meaningful for model recommendation.

| CPACS-Trans | 1D convolution is not used; |
| CPACS-Conv | Transformer is not used |
| CPACS-Add | In CPACS, vector addition is used instead of ACT-Skip () information fusion method. |
| CPACS-Cat | In CPACS, vector stitching is used instead of ACT-Skip () information fusion method. |
| CPACS-Mul | In CPACS, vector multiplication is used instead of ACT-Skip () information fusion method. |

Table 4 describes the comparison between the model variants.
### Table 2: Comparison of the Results of Different Forecasting Methods

<table>
<thead>
<tr>
<th>Model</th>
<th>Categorical Accuracy</th>
<th>Top5 Accuracy</th>
<th>Top10 Accuracy</th>
<th>Top15 Accuracy</th>
<th>Micro F1</th>
<th>AUROC</th>
</tr>
</thead>
<tbody>
<tr>
<td>GRU</td>
<td>0.1335</td>
<td>0.2281</td>
<td>0.2651</td>
<td>0.2833</td>
<td>0.1335</td>
<td>0.5853</td>
</tr>
<tr>
<td>GRU-Att</td>
<td>0.1726</td>
<td>0.269</td>
<td>0.3</td>
<td>0.3164</td>
<td>0.1726</td>
<td>0.6166</td>
</tr>
<tr>
<td>BiGRU</td>
<td>0.1441</td>
<td>0.232</td>
<td>0.2641</td>
<td>0.2786</td>
<td>0.1441</td>
<td>0.5932</td>
</tr>
<tr>
<td>Conv-GRU</td>
<td>0.2132</td>
<td>0.3238</td>
<td>0.3541</td>
<td>0.3733</td>
<td>0.2132</td>
<td>0.6489</td>
</tr>
<tr>
<td>Conv-BiGRU</td>
<td>0.2096</td>
<td>0.3199</td>
<td>0.3516</td>
<td>0.3705</td>
<td>0.2096</td>
<td>0.6449</td>
</tr>
<tr>
<td>Transformer-E</td>
<td>0.2189</td>
<td>0.3238</td>
<td>0.3534</td>
<td>0.3662</td>
<td>0.2189</td>
<td>0.6557</td>
</tr>
<tr>
<td>CPACS</td>
<td>0.2434</td>
<td>0.35</td>
<td>0.382</td>
<td>0.3994</td>
<td>0.2434</td>
<td>0.6739</td>
</tr>
</tbody>
</table>

### Figure 3: Influence of parameter changes

### Table 3: The variants of CPACS

<table>
<thead>
<tr>
<th>CAPCS Variants</th>
<th>Variants Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>CPACS-Trans</td>
<td>ACT-Skip([C_L, E_L]) , ACT-Skip([C_D, E_D])</td>
</tr>
<tr>
<td>CPACS-Conv</td>
<td>ACT-Skip([R_L, O_L]) , ACT-Skip([R_D, O_D])</td>
</tr>
<tr>
<td>CPACS-Add</td>
<td>Addition () Replace ACT-Skip ()</td>
</tr>
<tr>
<td>CPACS-Cat</td>
<td>Concatenate() Replace ACT-Skip ()</td>
</tr>
<tr>
<td>CPACS-Mul</td>
<td>Multiply() Replace ACT-Skip ()</td>
</tr>
</tbody>
</table>

### 5 CONCLUSION

This paper proposes a Customer Preference-Aware Customer Service Solutions Recommendation (CPACS) model. The model first uses the customer’s dynamic preference information to enhance and enrich the title information and description information of work order data, and then uses 1D-CNN and Transformer to capture the local and global timing dependence information in the work order data. In addition, an Additive Conv-Transformer Skip (ACT-Skip) method is used to fully integrate local and global dependent information to improve the recommended performance of the solutions.
Table 4: Performance comparison of different variants

<table>
<thead>
<tr>
<th>CAPCS Variants</th>
<th>Categorical Accuracy</th>
<th>Top5 Accuracy</th>
<th>Top10 Accuracy</th>
<th>Top15 Accuracy</th>
<th>Micro F1</th>
<th>AUROC</th>
</tr>
</thead>
<tbody>
<tr>
<td>CPACS-Trans</td>
<td>0.2347</td>
<td>0.334</td>
<td>0.3641</td>
<td>0.3806</td>
<td>0.2347</td>
<td>0.6594</td>
</tr>
<tr>
<td>CPACS-Conv</td>
<td>0.2291</td>
<td>0.3347</td>
<td>0.3673</td>
<td>0.3865</td>
<td>0.2291</td>
<td>0.6616</td>
</tr>
<tr>
<td>CPACS-Add</td>
<td>0.227</td>
<td>0.3333</td>
<td>0.3662</td>
<td>0.3847</td>
<td>0.227</td>
<td>0.6616</td>
</tr>
<tr>
<td>CPACS-Cat</td>
<td>0.2246</td>
<td>0.3368</td>
<td>0.3669</td>
<td>0.3823</td>
<td>0.2246</td>
<td>0.6629</td>
</tr>
<tr>
<td>CPACS-Mul</td>
<td>0.2225</td>
<td>0.3312</td>
<td>0.3582</td>
<td>0.3753</td>
<td>0.2225</td>
<td>0.6559</td>
</tr>
<tr>
<td>CPACS</td>
<td>0.2434</td>
<td>0.35</td>
<td>0.382</td>
<td>0.3994</td>
<td>0.2434</td>
<td>0.6739</td>
</tr>
</tbody>
</table>

ACKNOWLEDGMENTS

This work is supported by the Science and Technology Project of State Grid Shandong Electric Power Company: "Research and Application of Key Technologies of Big Data Analysis in Customer Service" (Grand No.2021A-102).

REFERENCES


