说　明　书　附　图

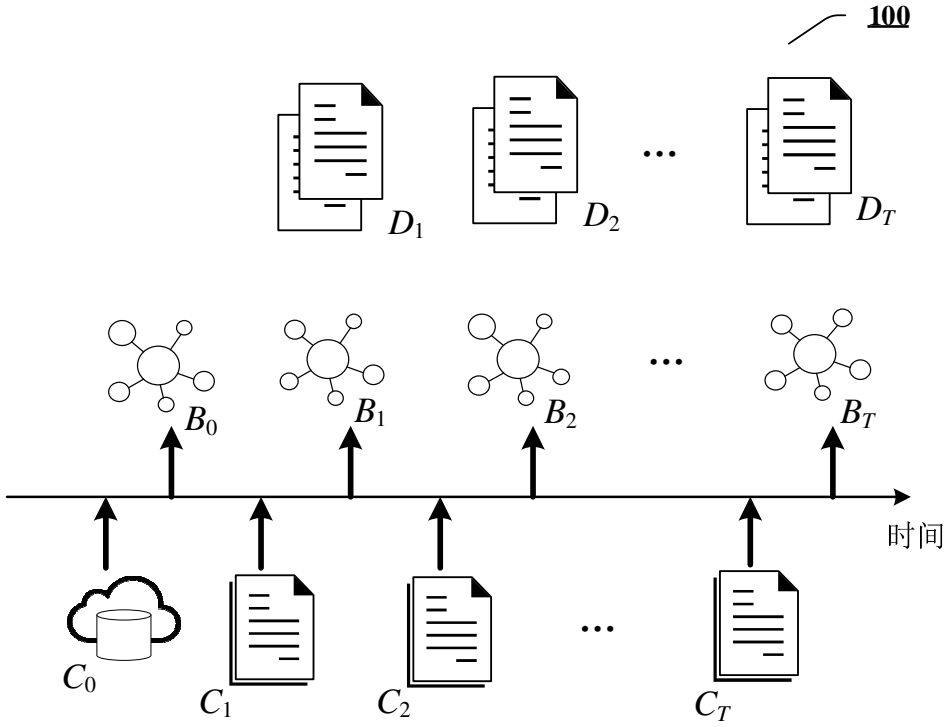
100
D_1
D_2
D_T
B_0
B_1
B_2
B_T
时间
C_0
C_1
C_2
C_T

图 1

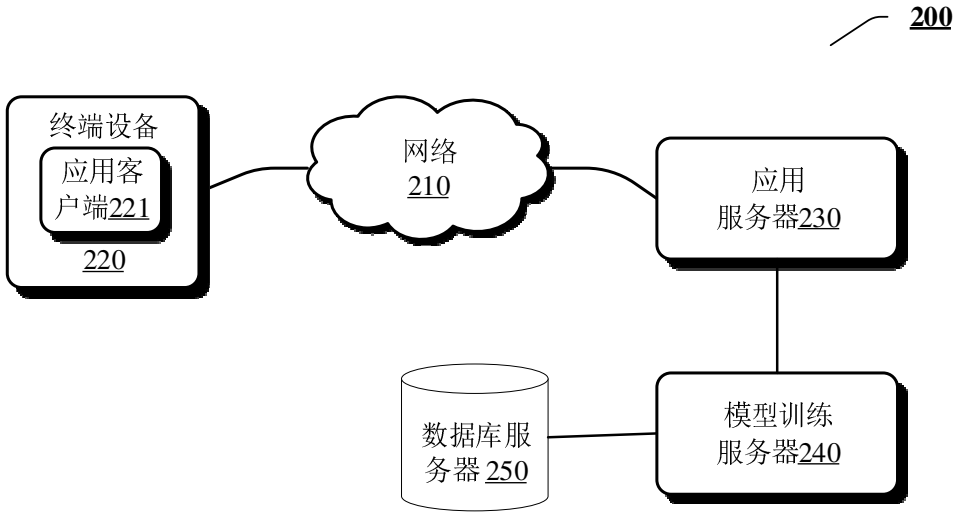
200
终端设备
应用客户端221
220
网络
210
应用服务器230
数据库服务器 250
模型训练服务器240

图 2

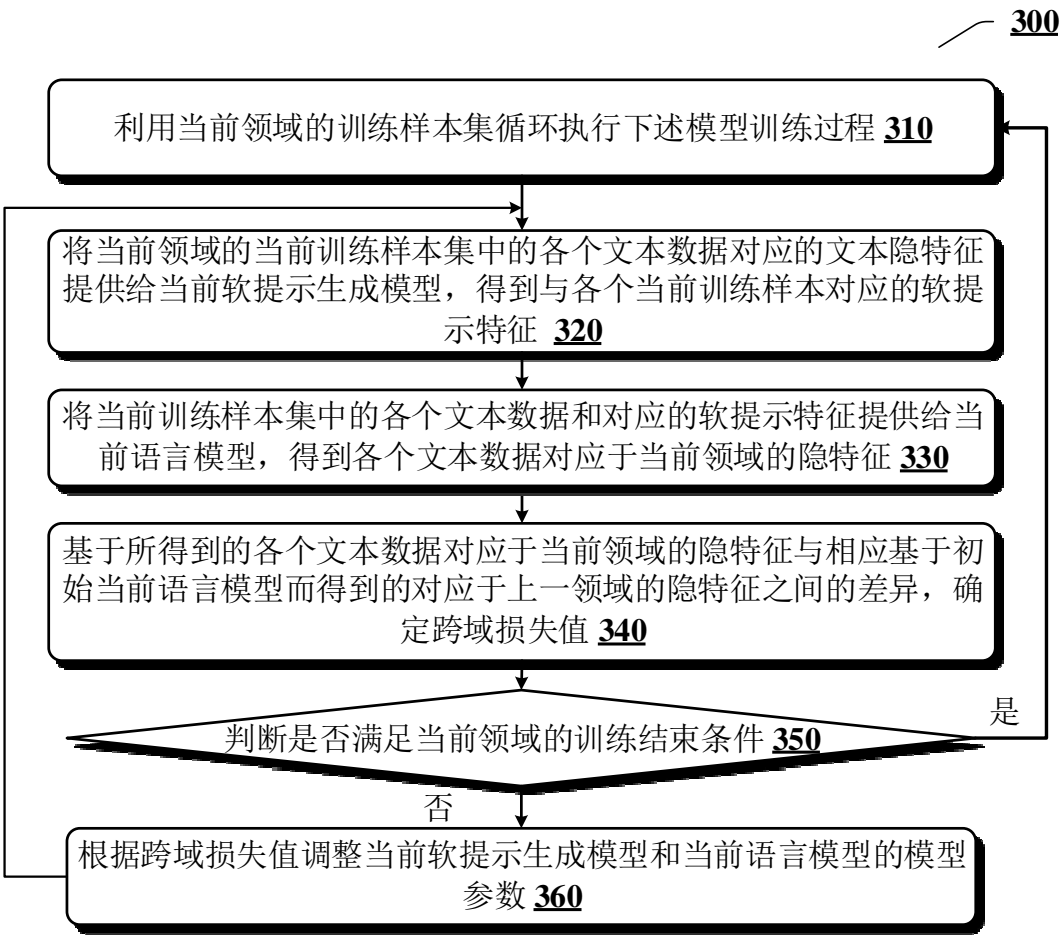
300
利用当前领域的训练样本集循环执行下述模型训练过程 310
将当前领域的当前训练样本集中的各个文本数据对应的文本隐特征提供给当前软提示生成模型，得到与各个当前训练样本对应的软提示特征 320
将当前训练样本集中的各个文本数据和对应的软提示特征提供给当前语言模型，得到各个文本数据对应于当前领域的隐特征 330
基于所得到的各个文本数据对应于当前领域的隐特征与相应基于初始当前语言模型而得到的对应于上一领域的隐特征之间的差异，确定跨域损失值 340
判断是否满足当前领域的训练结束条件 350
是
否
根据跨域损失值调整当前软提示生成模型和当前语言模型的模型参数 360

图 3

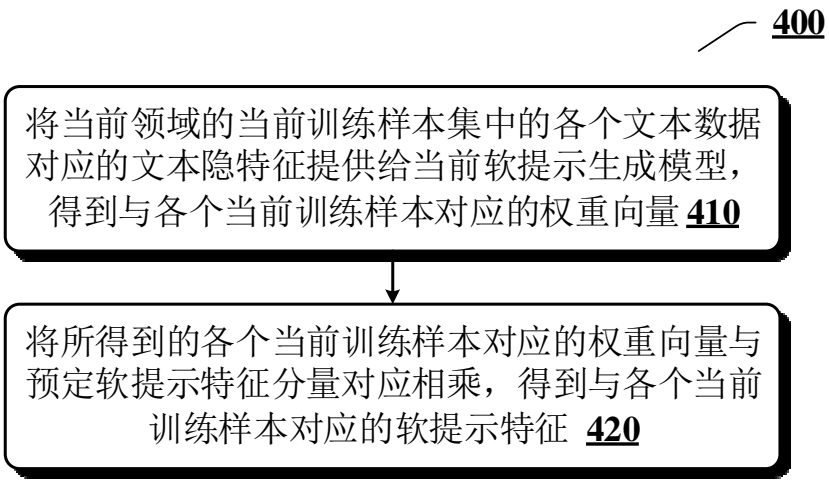
400
将当前领域的当前训练样本集中的各个文本数据对应的文本隐特征提供给当前软提示生成模型，得到与各个当前训练样本对应的权重向量 410
将所得到的各个当前训练样本对应的权重向量与预定软提示特征分量对应相乘，得到与各个当前训练样本对应的软提示特征 420

图 4

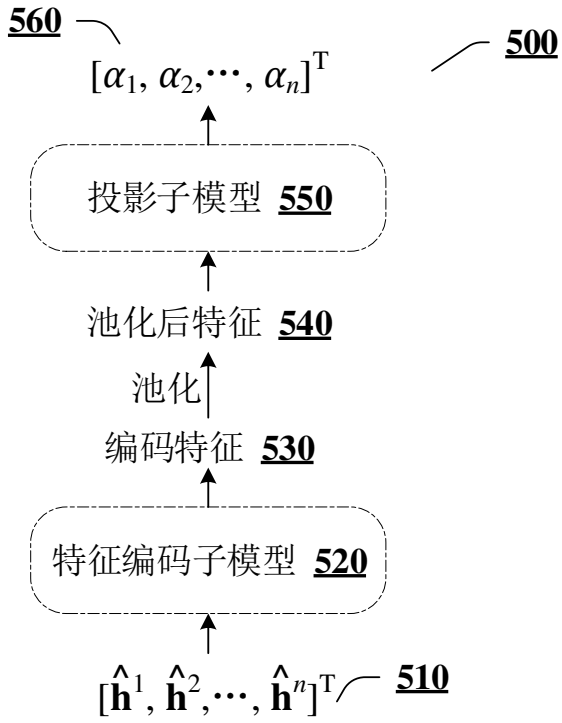
560
$[\alpha_1, \alpha_2, \cdots, \alpha_n]^T$
500
投影子模型 550
池化后特征 540
池化
编码特征 530
特征编码子模型 520
$[\hat{\mathbf{h}}^1, \hat{\mathbf{h}}^2, \cdots, \hat{\mathbf{h}}^n]^T$
510

图 5

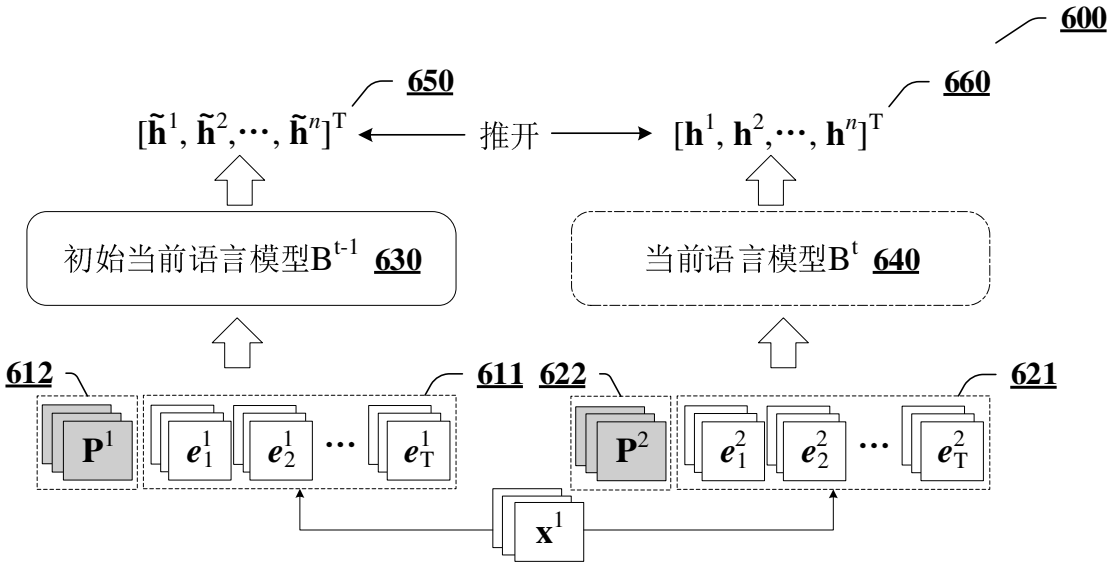
600
650
$[\tilde{\mathbf{h}}^1, \tilde{\mathbf{h}}^2, \cdots, \tilde{\mathbf{h}}^n]^T$
推开
660
$[\mathbf{h}^1, \mathbf{h}^2, \cdots, \mathbf{h}^n]^T$
初始当前语言模型B^{t-1} 630
当前语言模型B^t 640
612
611
622
621
$\mathbf{P}^1$
e_1^1 e_2^1 ⋯ e_T^1
$\mathbf{P}^2$
e_1^2 e_2^2 ⋯ e_T^2
$\mathbf{x}^1$

图 6

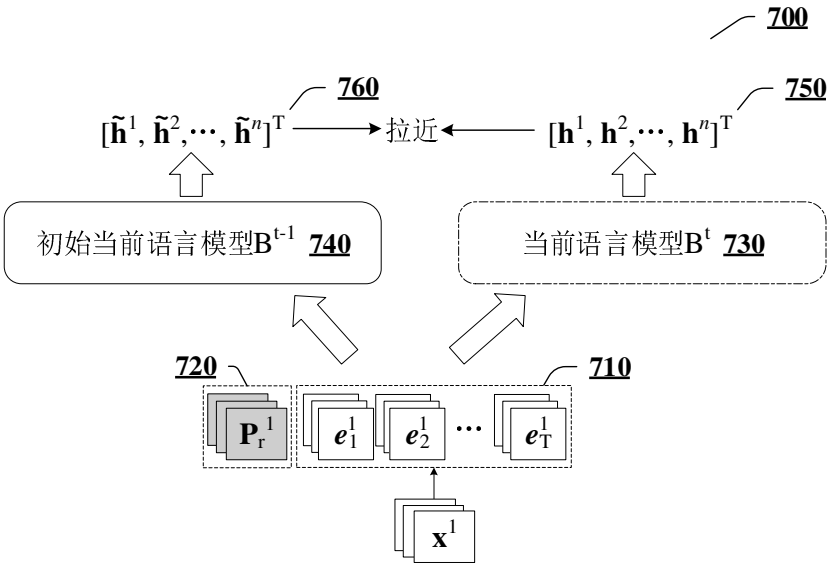
700
760
750
$[\tilde{\mathbf{h}}^1, \tilde{\mathbf{h}}^2, \cdots, \tilde{\mathbf{h}}^n]^T$
拉近
$[\mathbf{h}^1, \mathbf{h}^2, \cdots, \mathbf{h}^n]^T$
初始当前语言模型B^{t-1} 740
当前语言模型B^t 730
720
710
$\mathbf{P}_r^1$
e_1^1
e_2^1
e_T^1
$\mathbf{x}^1$

图 7

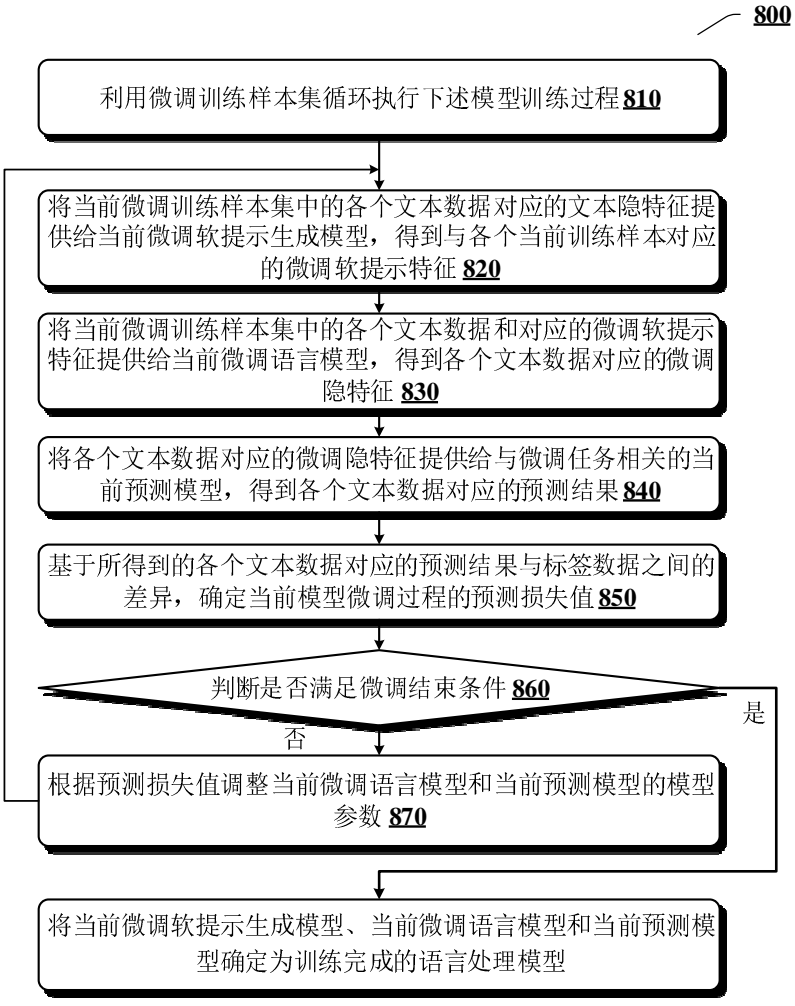
800
利用微调训练样本集循环执行下述模型训练过程 810
将当前微调训练样本集中的各个文本数据对应的文本隐特征提供给当前微调软提示生成模型，得到与各个当前训练样本对应的微调软提示特征 820
将当前微调训练样本集中的各个文本数据和对应的微调软提示特征提供给当前微调语言模型，得到各个文本数据对应的微调隐特征 830
将各个文本数据对应的微调隐特征提供给与微调任务相关的当前预测模型，得到各个文本数据对应的预测结果 840
基于所得到的各个文本数据对应的预测结果与标签数据之间的差异，确定当前模型微调过程的预测损失值 850
判断是否满足微调结束条件 860
是
否
根据预测损失值调整当前微调语言模型和当前预测模型的模型参数 870
将当前微调软提示生成模型、当前微调语言模型和当前预测模型确定为训练完成的语言处理模型

图 8


软提示生成模块 911
隐特征生成模块 912
损失确定模块 913
训练单元 910
参数调整单元 920
基于持续预训练的语言模型训练装置 900

图 9


微调特征生成模块 1011
预测损失确定模块 1012
训练单元 1010
参数调整单元 1020
语言处理模型的微调装置 1000

图 10

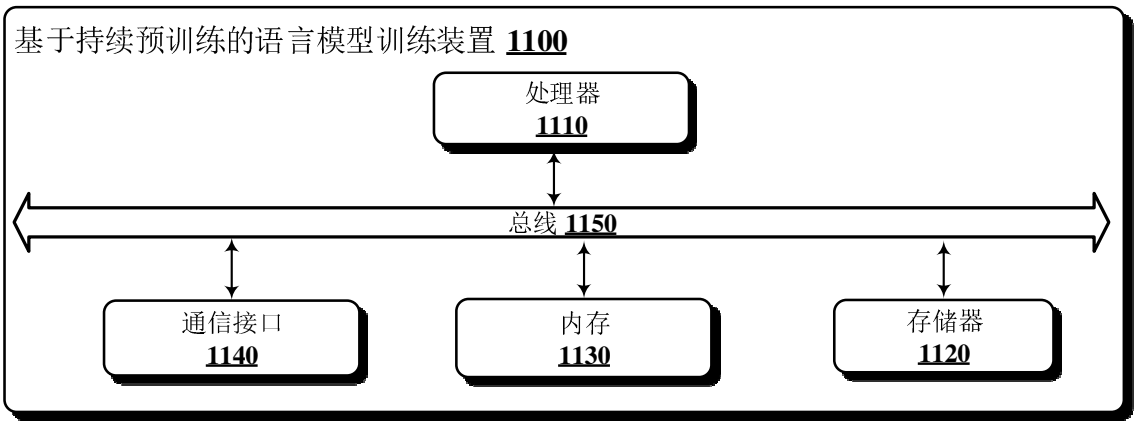
基于持续预训练的语言模型训练装置 1100
处理器
1110
总线 1150
通信接口
1140
内存
1130
存储器
1120

图 11

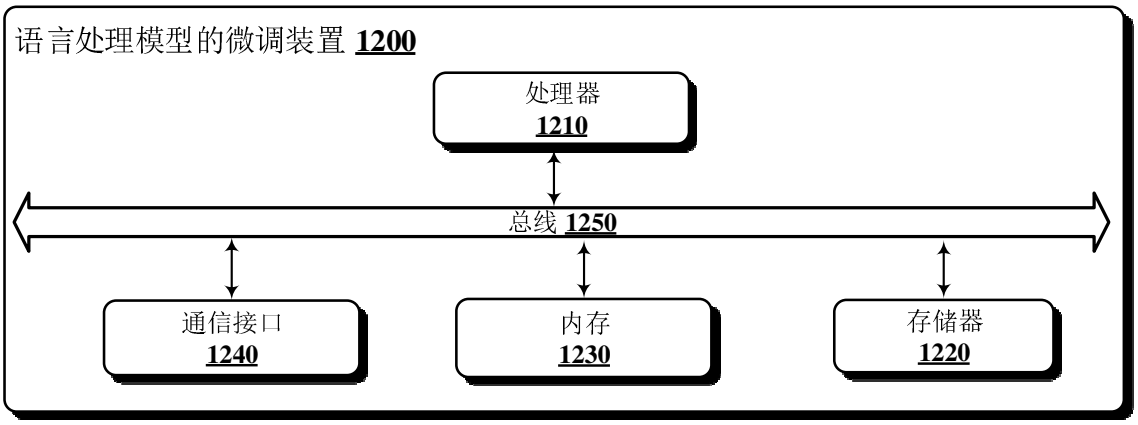
语言处理模型的微调装置 1200
处理器
1210
总线 1250
通信接口
1240
内存
1230
存储器
1220

图 12