用于抑郁症检测的机器学习算法及其比较。

Danish Muzafar[1], Furqan Yaqub Khan[2], Mubashir Qayoom[31]克什米尔中央大学信息技术系，查谟和克什米尔，印度[2]巴特那印度理工学院计算机科学与工程系[3]克什米尔大学物理系，查谟和克什米尔，印度 电子邮箱：furkaan309@gmail.com

摘要 –文本情绪智能在社交媒体平台上充分发挥人类情绪的作用，具有普遍的重要性。社交媒体平台富含情感内容，可用于观点挖掘、情绪挖掘和情感分析等多种目的。此类数据分析还可用于预防网络欺凌、自杀预防，以及检测社交媒体用户的抑郁症。本文通过分析在线社交媒体用户的行为，设计了一种自动抑郁症检测方法。所设计的抑郁症检测分类方法可有效挖掘用户的社交媒体互动，并判断用户是否患有抑郁症。该分类器采用情感人工智能领域的先进技术构建，包括长短期记忆网络（LSTM）及其他机器学习分类器。长短期记忆网络（LSTM）分类器的最高准确率约为70%，而支持向量机（SVM）的最高准确率为81.79%。我们使用文献中广泛用于情绪挖掘任务的数据集对分类器进行了训练。此外还给出了结果的混淆矩阵。

关键词——情绪、人工智能、深度学习、机器学习算法、词嵌入、推特。

一、引言

抑郁症会引发悲伤、沮丧、缺乏兴趣和烦躁。它会导致各种情绪和身体问题，并可能削弱一个人充分发挥潜能的能力 [1]。人类的心理状态问题极其严重。在最糟糕的情况下，抑郁症是导致8万人死亡并走向自杀的主要原因。青少年最容易患上抑郁症，因为他们往往需要心理治疗，

青少年无限制地使用手机和社交媒体，在现实世界中花费的时间更少，而在虚拟世界中花费的时间更多，这导致他们的内心状态与外部世界之间产生鸿沟，进而造成强烈的孤独感，并最终导致抑郁症[2]。基于社会地位、语言水平、种族、族裔、移民标准、健康、行为伦理等因素，通过社交媒体实施的霸凌，是社交媒体导致用户抑郁的主要弊端。[4] 本研究论文所要解决的问题是确定研究对象是否具有抑郁症症状。借助神经网络，我们利用用户的推文来诊断研究对象的行为是积极还是消极。如果一个人的推文表现出一定程度的积极性，我们就可以预测其未患抑郁症；否则，则说明其患有抑郁症。本文采用的抑郁症判定标准是：如果一个人的推文中有80%被归类为负面内容，则认为其处于抑郁状态。我们通过基于科学认可的80%标准的先进机器学习和深度学习技术，分析用户在推特账号上发布的信息流。[6][7]。

1.1. 抑郁症的一些初始症状如下：

- 。 感到疯狂、悲伤或情绪低落。 体重减轻。 疲劳加重和睡眠困难。。 精力丧失、疲劳加重，并有自杀冲动。。思考十分困难，难以注意力集中或作出决定，并以对任何事物都不感兴趣为前提进行思考。。 频繁且持续地想到死亡或自杀

症状必须持续至少两周或更长时间。抑郁症对不同性别的人可能产生不同影响；研究表明，患有抑郁症的男性可能出现易怒、逃避或危险行为，或者表现得极其愤怒。要诊断为抑郁症，症状必须持续至少两周甚至更长时间 [1]

1.2. 抑郁症的风险因素- 即使一个人看起来生活在相当理想的条件下，抑郁症也可能

抑郁症可能由多种情况引起：

- 。生物化学：在某些情况下，激素差异可能会导致大脑中的化学物质失衡，从而

- 遗传学：遗传学在共同抑郁中也发挥着重要作用。例如，如果一对同卵双胞胎

1.3 社交媒体与抑郁症的作用

社交网络正日益成为人们生活中重要的一部分，它们映照出用户在社交媒体上的个

II. 抑郁症与情绪智力

如果一个人与他人关系亲密，那么可以说这个人对对方有情感依恋；如果这种情感

或所有人患上抑郁症

2.1. 情绪智力通常由以下四个属性来定义：

a. 自我管理 – 你能够控制冲动的情绪和行为，以健康的方式处理情绪，采取行

b. 自我意识 – 你能识别自己的感受，以及自己的想法和行为如何受到周围环境

c. 社会意识 – 你具备同理心。你应该能够理解他人的想法、愿望和担忧，察觉

d. 关系管理 – 你知道如何建立并维持良好的人际关系，清晰地沟通，激励并影

III. 讨论

3.1. 数据集

我们通过手动标注推特数据，构建了一个抑郁症检测数据集。本研究采用两个类别：

本文将逐一讨论所使用的不同算法：

3.2. 支持向量机（SVM）

支持向量机（SVM）是一种用于机器学习的监督学习算法，可用于分类和回归问题。不过，在分类问题中，它的用途最广 [1]。在 SVM 算法中，我们将每个数据项表示为 n 维空间中的一个点（其中 n 是特征的数量），每个

特征的值即为坐标值。接着，我们通过确定能够清晰区分两个类别的超平面来进行分类 [12]。支持向量机是一种判别式分类器，通常由一个唯一的超平面来描述（SVM）。换句话说，给定带标签训练数据，该算法会生成一个理想的超平面，用于对新样本进行分类（监督学习）。这个超平面是二维空间中的一条直线，将平面分成两个区域，每个类别位于其中一侧。在线性支持向量机中，通过运用一些线性代数转换问题，从而完成超平面的学习 [10]。此时，核函数便发挥作用。在我们的实验中，我们使用 80-20% 的训练测试集划分来训练线性支持向量机。我们使用 scikit-learn 的核函数进行分类。

3.3. 随机森林分类器

随机森林分类器是一种集成算法，它根据从训练集中随机选取的子集生成一组决策树。为了评估测试对象的最终类别，它会汇总各棵决策树的投票结果。不同的决策树可能生成彼此重叠的子集。为了解决这一问题，我们采用加权原则来考虑任一决策树结果的影响。错误率高的树赋予较低的权重，而错误率低的树则赋予较高的权重。因此，错误率较低的树对决策的影响更大。待构建的树的总数也是决策树的重要特征。最小分裂数、分裂准则等都是随机森林分类器的基本参数。我们使用了包含 40 个估计器的随机森林分类器。作为分类器输入的训练集和测试集，按总数据条目数的 80:20 划分为训练测试集。我们还使用了 scikit-learn 核函数进行分类。

3.4. 多项式朴素贝叶斯

朴素贝叶斯是一类基于贝叶斯定理的算法，它们采用强（朴素）假设，即每个特征都独立于其他特征，以此预测给定样本所属的类别。由于它们是概率分类器，因此会使用贝叶斯定理计算每个类别的可能性，并选择概率最高的类别。朴素贝叶斯分类器已成功应用于多个领域，包括自然语言处理（NLP）。处理自然语言处理问题时，也可以选择支持向量机（SVM）和神经网络等其他方法。

另一方面，朴素贝叶斯分类器的简单性使其特别适用于这类分类任务。此外，事实证明，它们在各种自然语言处理应用中都快速、可靠且准确。我们采用非朴素贝叶斯方法整体分析句子，因此，如果某个句子未出现在训练集中，就会得到零概率，导致无法进行后续计算。然而，在朴素贝叶斯方法中，每个词都被假设为独立于其他词 [10]。现在，我们不再考察完整句子，而是将单个词作为短语来分析。

3.5 长短期记忆（长短期记忆网络）

长短期记忆网络已被证明在涉及序列知识的任务中十分有效。然而，其他分类器使用的输入向量（例如 TF-IDF 权重向量）并不保留文档中各术语和短语之间的序列关系信息。因此，长短期记忆网络更适合使用词嵌入向量序列作为输入；事实上，在现代深度学习中，词嵌入是文本表示的首选方式。

我们使用训练集和调参集调整了 LSTM 的记忆单元数量、训练轮次、批量大小，以及输入和循环丢弃率。训练轮次是指神经网络对整个训练集进行前向和反向传播的次数[12]。如果训练轮次过少，网络可能无法充分拟合，得到的神经元权重也可能远非最优。批量大小是指每次更新权重前输入网络的样本数量；它能降低内存需求和达到最优权重所需的迭代次数，从而提高训练效率。最后，为尽量减少过拟合，丢弃率指的是在更新权重时，某个输入节点或循环节点被排除在计算之外的概率。按照 80-20 训练测试划分，分别将输入提供给训练集和测试集。

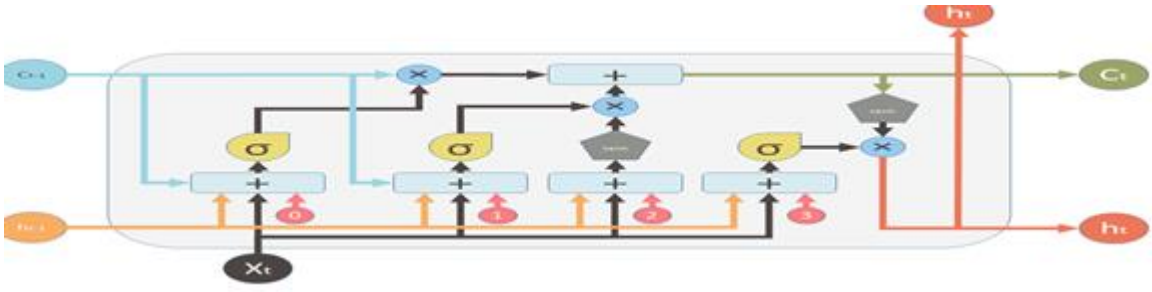

图1：LSTM 分类器的机制

标准的 LSTM单元由一个单元、一个输入门、一个输出门和一个遗忘门构成。单元

中的关键事件之间可能存在时长不确定的滞后。长短期记忆网络非常适合对时间序列数据进行分类、处理并据此作出预测。长短期记忆网络旨在解决训练常规循环神经网络时可能出现的梯度消失问题。与循环神经网络、隐马尔可夫模型及其他序列学习方法相比，长短期记忆网络对间隔长度的相对不敏感性，使其适用于多种应用。

A. LSTM架构：LSTM单元存在多种架构。标准架构由一个细胞（LSTM单元的记忆

入门、输出门和遗忘门。LSTM设备的其他版本可能没有其中一个或更多门，甚至还可能没有其他门，例如，门控循环单元（GRUs）没有输出门。输入门决定有多少新值进入细胞，遗忘门决定有多少细胞值得以保留，而输出门决定有多少细胞值用于生成LSTM单元的输出激活。LSTM各门的激活函数构件是逻辑Sigmoid函数。

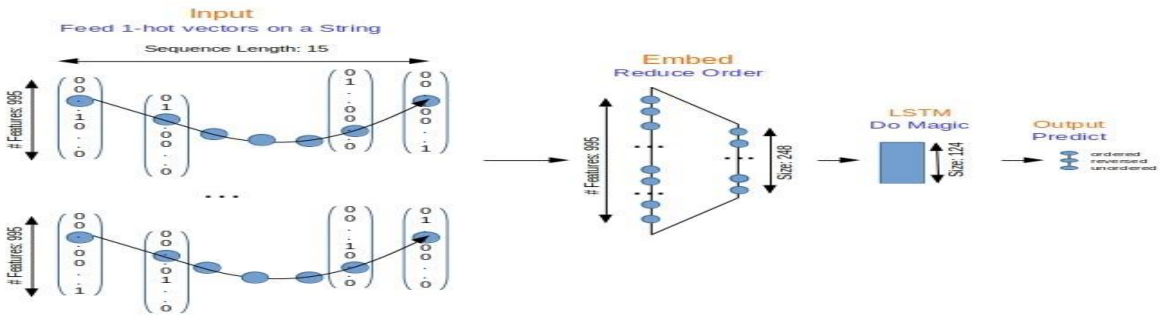

图2：LSTM的过程

B. 词嵌入

词嵌入是一组用于表示词语和文本的方法，采用稠密向量表示。与之不同的是，在

每个向量都反映了该词在连续向量空间中的投影 [2]。词语的位置取决于它在使用

的向量空间中的位置。从文本中学习词嵌入的两种常见方法包括：

- Word2Vec。
- GloVe。

本文使用了 Word2Vec 和 Doc2Vec。

C. Keras 嵌入层

Keras 为神经网络提供了一个可用于文本数据的嵌入层。该层包括对输入数据进行整数编码，使每个单词都由一个唯一的整数表示。Keras 同时提供的 Tokenizer API 可用于执行这一数据准备阶段。由于嵌入层使用随机权重进行初始化，训练数据集中的所有单词都会学习到一个嵌入向量。这是一个灵活的层，可以通过多种方式应用。例如，它可以单独使用来学习词嵌入，然后将其存储起来，并在其他模型中重复使用；也可以作为深度学习模型的一部分，使嵌入与模型同步学习。它属于迁移学习的一种，也可用于加载预训练词嵌入模型。嵌入层会学习权重。如果将模型保存为文件，嵌入层的权重也会一并保存。嵌入层的输出是输入词序列（输入文档）中的一个二维向量，其中每个单词对应一个嵌入向量。如果希望使用展平层将全连接层直接连接到嵌入层，则必须先将二维输出矩阵展平为一维向量。最后，我们训练长短期记忆网络分类器，用于训练数据和测试数据的分类。

IV. 方法论

本节讨论利用用户的在线社交资料进行抑郁症检测所采用的整体方法。我们采用情感人工智能执行抑郁症检测任务。我们使用深度学习框架和机器学习创建抑郁症分类器，并利用该分类器对来自在线社交

资料的推文进行自动抑郁症检测。图5展示了该任务采用的整体方法。

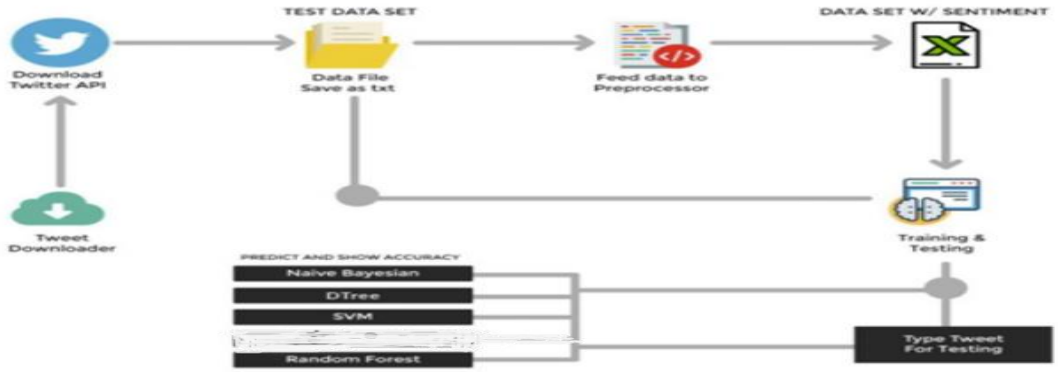

图3：抑郁检测系统的总体方法。

该系统接收一个情绪标注数据集作为输入，用于训练监督式深度学习分类器或机器学习分类器。分类器在情绪标注数据集上完成训练后，我们使用来自在线社交资料的推文进行抑郁症检测。如果某个资料中的推文有超过75%被分类为抑郁，则我们的系统会为该资料标记红旗，表示其可能存在抑郁症状。图6详细展示了该系统的各个组件。

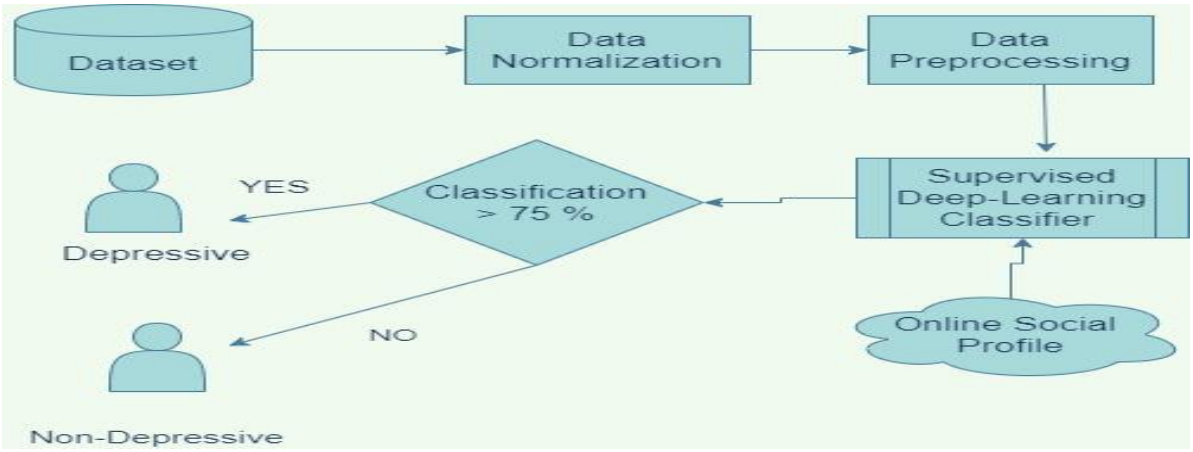

图4：系统架构

该系统包含 5 个模块，并使用情绪标注数据集。

4.1. 数据规范化

作为机器学习数据准备的一部分，标准化是一种有时会用到的技术。调整的目的是

此处，你还删除了所有转推、非字母数字字符、网址和 @提及。除此之外，除第一和第三人称代词外，所有停用词都已移除。大约 40% 的工作由自然语言工具包和所有库自动完成。

4.2. 数据预处理

在将每条推文传递给自动分类器之前进行预处理，是预处理模块为分类器做好准备的方式。通过使用英语目录，可以从推文中消除俚语和缩写。如果发现某个术语没有有意义的定义，则将所有词语传递到词典模块中进行查找，然后传递到词语替换模块中，用正确的词语替换。在词语替换模块中，我们采用了短信词典、Netlingo 和城市词典。预处理阶段由以下步骤组成：

4.3. 分词

使用斯坦福核心自然语言处理软件包执行该操作，将推文拆分为句子和词语，例如词语、句子或主题等单位？在本例中，提取包含推文的 CSV文件的第一列，并将其转换为独立的词元。

4.4. 词干提取

词干提取是将词语简化为最基本形式的过程。这样我们就能将相似的词语归为一组。此处使用 Poster词干提取器来实现。

4.5. 停用词移除

由于这些广泛使用的词语对训练几乎没有用处，因此需要将其省略；如果不加以忽略，还可能导致错误的结果。自然语言工具包库中有一个停用词集合，可作为从数据集中移除停用词的指南。使用词频-逆文档频率、斯坦福和维基特征从推文中删除的停用词。

4.6. 词性标注器

通过使用词性标注器将分词后的文本分配到相应的词性部分，以提高剩余数据的质量。由于其他词性并不十分重要，因此可以仅移除形容词、名词和副词。例如：从“I LOVE CODING”中提取出“I”，“I”是代词，其余部分被省略。

4.7. 词形还原

该方法用于将词语还原为其词干。词形还原是指借助词汇表和词形分析，按照规则处理词语，主要通过移除屈折词尾，将词语还原为其根形或词典形式，即词元；此处使用的是 Stanford CoreNLP 工具包。

完成所有这些预处理步骤后，就会生成词袋。词袋会统计每个词的出现次数，随后

4.8. 词袋或独热编码

在这种方法中，向量的每个部分都对应语料库词汇表中的一个词或 n 元语法（词元）。

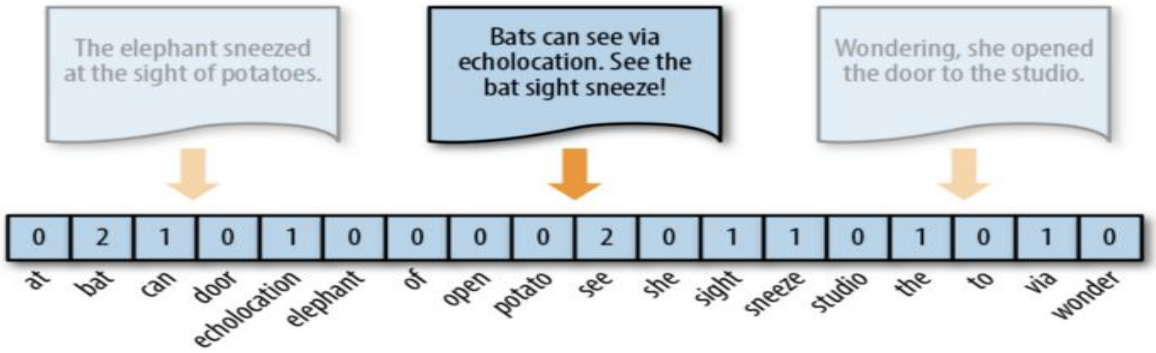

图 5：词袋表示

在上述示例中，我们的语料库由三个文档中出现的所有不重复词语组成。词袋模型，

可用于以多种方式从文档中提取信息。词袋是对文档中出现词语的一种文本表示。

它涉及两个方面：

1. 已知词汇的词汇表。

2. 应考虑已知词汇的出现情况。

之所以称为词袋，是因为文本中有关词语顺序或排列的所有信息都会被舍弃。该模

分析文本中各词语的表示图。词袋方法（BOW）将每个词频视为一个特征。是一种常用于句子和文本的角色提取方法。其假设是，内容相近的文档也彼此相似。此外我们仅凭文档内容就能推断其重要性。词袋可以构建得很简单，也可以很复杂。难点在于如何创建已知词汇（或词元）的词典，以及如何评估这些词汇的出现情况。

4.9 词频-逆文档频率

TF-IDF 是词频-逆文档频率的缩写。它是一种统计指标，用于确定某个词对于文档语料库中的文本有多重要。这种重要性与该词在文本中出现的次数有关，但也会根据语料库中包含该词的文档数量进行调整。

4.10 计数向量化

可以使用计数向量化统计每个词在文档（即独立文本，例如文章、书籍，甚至只是

CountVectorizer 的工具，可以帮助完成这项工作。例如：“今天天气很好，所以我到户外去享受这美好而晴朗的天气。”在下面的统计结果中，“the”、“weather”和“and”都出现了两次，而其他词只出现了一次。计数向量化正是通过这种方式实现

4.11 基于分布相似性的表示- **词嵌入**

词嵌入是一种学习得到的文本表示（实数向量），在这种表示中，含义相关的词会[illegible]于每个词，采用稠密的分布式表示这一理念是该方法的核心。每个词都可以表示为一个包含几十甚至数百个维度的实值向量。相比之下，独热编码等分布式词表示需要数十甚至数百万个维度。最流行的两种词嵌入是 Word2Vec 和 Glove：

4.12. Word2Vec：

Word2Vec 主要有两个版本——Skip-Gram 和连续词袋（CBOW）。CBOW 模型[illegible]Skip-Gram 模型则通过预测周围的词（上下文）来学习。顾名思义，Word2Vec 用一组特定的数字（称为向量）表示每个单独的词。向量的基本函数是计算向量之间的余弦相似度，用以表示这些向量所对应词语之间的语义相似程度。Skip-Gram 旨在预测给定词的直接邻近词。我们选取一个中心词和一个上下文窗口（邻近词），[illegible]

[illegible]

[illegible]过最小化损失函数，预测模型便能学习这些向量。在 Word2Vec 中，这一过程通过执行语言建模任务（预测下一个词）的前馈神经网络实现，并使用随机梯度下降等优化技术。这些特征可用于我采用的机器学习算法，例如支持向量机、朴素托马斯贝叶斯和随机森林。

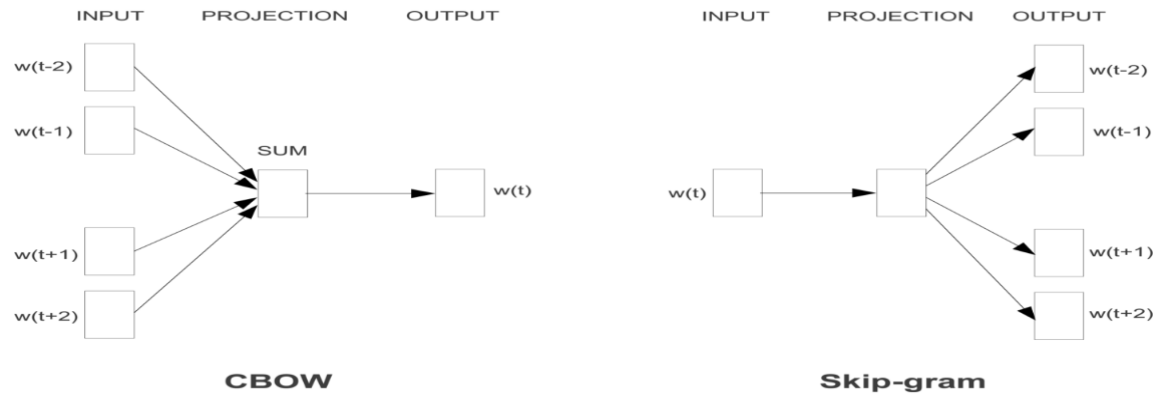

图 6：Word2Vec CBOW 与 Skip-gram

4.13. 用于长短期记忆网络的 Gensim Doc2Vec 模型

Gensim 被视为一个自然语言处理工具包，旨在实现“面向人类的主题建模”。但从

Doc2Vec，这属于深度学习算法模型。因此先来说明 Doc2Vec，我已经提到过在机器学习算法中使用的 Word2Vec。因为我既使用了机器学习算法，也使用了深度学习算法。因此，在机器学习中，我使用了 Word2Vec；在 LSTM 模型中，我使用了 Doc2Vec。正如前文所述，Doc2Vec 的目标是构建文本的数值表示，不论文本长度如何。但是，记录下你的语言并只有这种形式，因此需要找到另一种技术。由于

是一个有属于标签（主题的）的向量吗？仅仅想想，我们还可以还与不必是有唯一性的向量；例如，如果我们的文档带有标签（确定如此），就可以将这些标签添加进去，并将其表示为向量。

此外，这些标签不必是唯一的。这样，我们就可以在唯一的文档标签之外添加 17

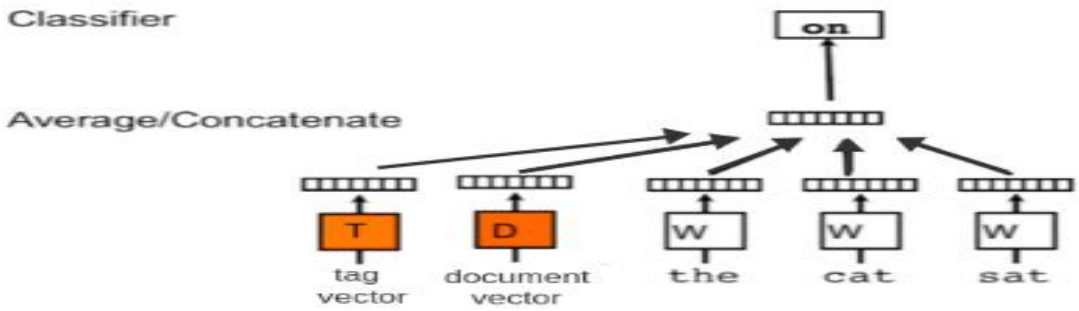

图 8：带有标签向量的 Doc2Vec 模型

4.14. 训练

分类器需要训练集中的标签和推文。在此示例中，训练集指的是必须先经过处理、

4.15. 保存分类器和计数向量化器对象

由于训练只需进行一次，因此有必要将训练好的分类器对象加载到 pickle 文件中。

4.16. 测试

已保存模型的加载：从 pickle 文件中加载符合要求的分类模型，用于测试数据集预

对测试推文进行类别预测。每条推文都会被划分为抑郁或非抑郁类别。我们根据真

正例、假正例、真阳性和假阳性的值计算混淆矩阵。以评估分类结果。混淆矩阵用

于衡量分类器给出准确结果和错误结果的情况。

V. 结果

我们设计的抑郁检测系统使用情感人工智能，通过在线社交资料进行抑郁症检测。我们使用了监督学习模型

和深度学习算法，即长短期记忆（LSTM），并使用情绪数据集对分类器进行训练。该数据集包含 15,000 条由我们手动标注的推文。分类器的最高准确率约为 70%（长短期记忆网络），而支持向量机的最高准确率为 81.79%。我使用文献中广泛用于情绪挖掘任务的数据集对分类器进行了训练。结果通过准确率和混淆矩阵等分类指标呈现。

混淆矩阵： 混淆矩阵是一种机器学习分类性能评估方法。它是一种表格，可用于查

例（TP）、真负例（TN）、假正例（FP）和假负例（FN）。用于计算准确率、召回率、精确率和F1分数。借助这四个类别，我们可以计算所有数值。评估依据是F1分数和准确率。F1分数是主要性能指标，准确率是次要指标。借助混淆矩阵，我们可以计算精确率、召回率、F1分数和准确率；这些指标简单地展示了分类器结果的准确程度，计算公式如下：

A. 精确率：TP/TP+FP，其中TP为真正例，FP为假正例。

B. 召回率：TP/TP+FN。

C.FI 分数：2*（精确率*召回率）/ 精确率 + 召回率

该数据集经过平衡，抑郁类和非抑郁类各包含 7,500 条推文。支持向量机分类器的准确率高于其他分类器。我们选择支持向量机分类器处理未见数据。我们使用长短期记忆网络（深度学习分类器）、支持向量机、朴素贝叶斯和随机森林分类器（机器学习分类器）训练了系统。各分类器的准确率如下表所示：

分类器	精确率	召回率	F1分数	准确率
LSTM	0.88	0.60	0.71	70.00%
多项式朴素贝叶斯	0.79	0.78	0.78	79.32%
SVM	0.79	0.82	0.79	81.79%
随机森林	0.75	0.82	0.78	80.37%

表1 不同分类器的性能结果

图9、10、11和12展示了系统使用混淆矩阵的性能：

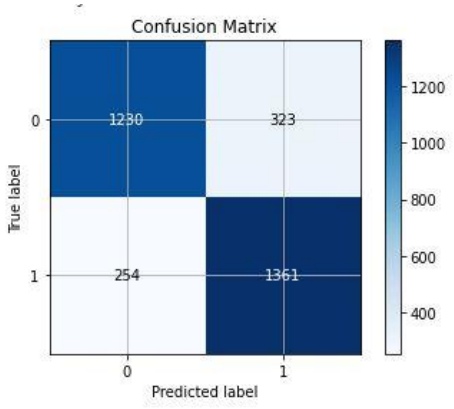

图9：支持向量机的混淆矩阵

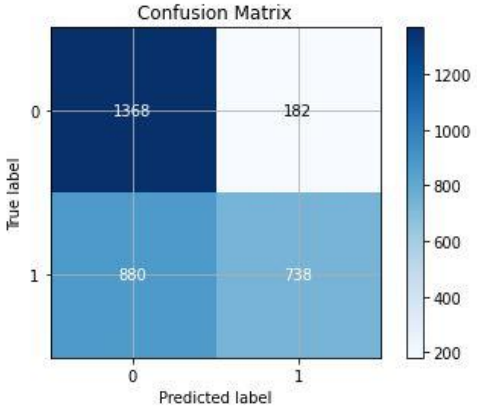

图10：长短期记忆网络的混淆矩阵

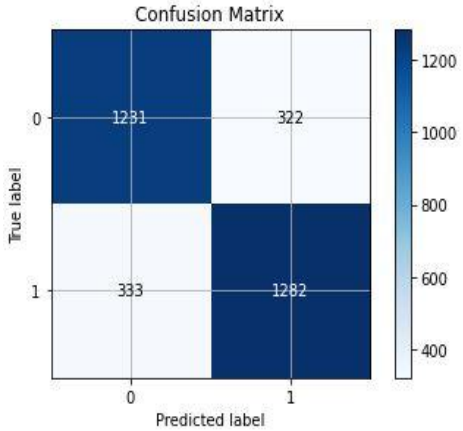

图10： 多项式朴素贝叶斯

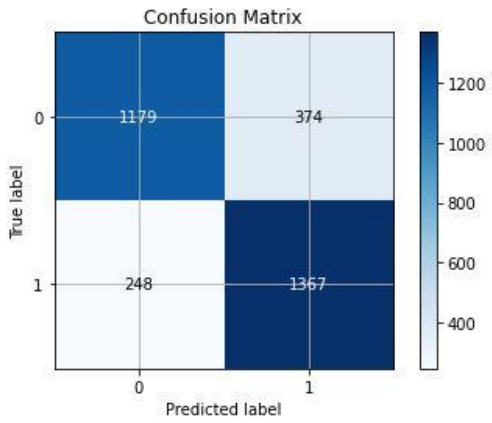

图 12：随机森林

V1. 结论与未来工作

对所选数据集进行的分析为探讨相关问题提供了一些见解。以下概述了我们的研究

非你或情绪低落，但有些人会经历这些感受，持续时间

每天持续较长时间（数周、数月，甚至数年），有时也没有明显原因。沮丧不仅仅是一时的情绪，而是一种个人的个人生活方式的体现，任何人都可能在任何时候患上抑郁症。然而，现代社会中的某些阶段或外接会使人更容易产生焦虑和抑郁症。在人的一生中，个体周围或内在的变化都会引发情绪波动，这可能使性格内向、社交孤立的人感到不快乐。我们比较了当前最先进的深度学习模型，以便从用户层面的推文中提前检测抑郁症。我们使用人工预处理的数据集运行模型，发现支持向量机表现得了更好的结果。未来，我们可以采用一种文件的方法，从更广泛的情绪特征

参考文献

[1] Mander Deshpande 和 Vignesh Rao，《利用情绪人工智能》检测抑郁症，IEEE，2017年。

[2] Marcel Trotzek、Sven Koitka 和 Christoph M. Friedrich，《利用神经网络和语言学元数据早期检测文本序列中的抑郁迹象》，IEEE，2018年。

[3] Adyan Marendra Ramadhani 和 Hong Soon Goo，《使用深度学习方法进行推特情感分析》， 7th 国际年度工程研讨会（InAES），印度尼西亚日惹，2017年。

[4] Wei Tong Mok、Rachael Sing、Xiuting Jiang 和 Swee See，《社交媒体与抑郁症研究》，IEEE，2014年。

[5] Niko Colneric 和 Janez Demsar，《推特上的情绪识别：比较研究与统一模型训练》，IEEE，2018年。

[6] Mohammad Jabreel 和 Antonio Moreno，《基于深度学习的推文多标签情绪分类方法》，应用科学，MDPI，2019年。

[7] Hakak Nida、Kirmani Mahira、Mohd. Mudasir、Muttoo Mudasir Ahmad 和 Mohd. Mohsin，《自动情绪分类器》，Springer，2019年。

[8] Krishna Shrestha, 《使用 Twitter 数据进行抑郁症诊断的机器学习》，使用 Twitter 数据，《国际计算机工程研究趋势期刊》（IJCERT），2018年。

[9] Nida Manzoor Hakak、Mohsin Mohd、Mahira Kirmani 和 Mudasir Mohd，《情感分析：一项调查》，IEEE，2017年。

[10] Hemanthkumar M、Latha，《基于推文情感分析的抑郁症检测》，《国际工程与技术研究期刊》（IRJET），2019年。

[11] Md. Rafiqul Islam、Muhammad Ashad Kabir、Ashir Ahmed、Abu Raihan M. Kamal、Hua Wang 和 Anwaar Ulhaq，《使用机器学习技术从社交网络数据中检测抑郁症》，Springer，2018年。

[12] Aswathy K S、Rafeeque P C 和 Reena Murali，《利用推特中的深度学习方法检测抑郁症》，国际系统、能源与环境会议，2019年。