# Image and Attribute Based Convolutional Neural Network Inference Attacks in Social Networks

# 摘要

在现代社会中，社交网络对于在线用户起着重要作用。然而，服务兴旺背后的一个不可忽视的问题是隐私问题。同时，近年来神经网络得到了迅速发展，并被证明在推理攻击中非常有效。本文提出了一种新的社交网络推理攻击框架，它巧妙地集成和修改了现有的最先进的卷积神经网络（CNN）模型。因此，无论用户是否拥有合法的配置文件映像，该框架都可以适用于推理攻击的更广泛的适用场景。此外，该框架能够提高现有的高精度CNN以用于敏感信息预测。除了框架之外，本文还展示了用于推理攻击的完全连接神经网络（FCNN）的详细配置。这部分通常在现有的研究中不存在。此外，实施传统的机器学习算法来比较来自构建的FCNN的结果。本文还讨论了应用差分隐私（DP）可以有效地破坏社交网络中推理攻击的准确性。

关键词：推理攻击，社交网络，神经网络，机器学习

# 简介

现在，在线社交网络是每个人的重要组成部分。 在美国，人们花在网络社交网络上的时间不断增加; 网上花费的全部时间中有30％现在分配给社交网络互动。 社交网络平台不断发展他们的工具和选项以吸引新的观众

在线社交网络每天都会发送数十亿和数十亿条消息。 然而，服务兴旺背后的一个不可忽视的问题是隐私问题。 当一个人为社交网络注册一个有效的帐户时，必须创建一个配置文件，以确保他的家人，朋友和同事能够识别他自己。 该配置文件包含多条信息。 其中有些是强制性的，有些是可选的。 根据用户的喜好，他需要在表达自己的足够信息和隐藏敏感个人信息之间找到平衡点。 然而，通过将卷积神经网络（CNN），全连接神经网络（FCNN）或其他机器学习算法发起的攻击部署到大量可用数据上，可以推断隐藏的敏感个人信息，并且用户的隐私会受到影响。

本文的动机在于敏感信息攻击对在线用户和广告商都有深远的影响。 了解这些攻击可以从根本上帮助制定防御措施，防止在线用户泄露隐私。 用户的隐私暴露出来会导致一些后果。 首先，由于当前密码恢复机制在发送密码恢复链接之前通常会询问用户的敏感信息，因此敏感信息可能有助于攻击者恢复用户的密码。 其次，在线用户隐私泄露可能会影响用户的离线活动。 例如，了解用户的详细信息，例如姓名，生日和地址可以帮助伪造信用卡甚至身份证件。 第三，敏感信息可以帮助广告商为目标用户提供广告。 例如，了解用户的年龄，性别和邮政编码可以揭示用户的生活方式，从而大大提高广告投放的准确性。

在本文中，进行了广泛的研究，以从用户的个人资料图像和不敏感属性中推断敏感的个人信息。 具体而言，新浪微博数据用于根据可用的公共用户配置文件推断敏感年龄信息。 新浪微博是中国最受欢迎的社交网络。 每个微博资料都包含一个个人资料图片和许多属性，如位置，性别，工作信息，关注次数，帖子数量等等。 在这些属性中，敏感的年龄信息通常是隐藏的。 因此，本文重点根据他可用的公开个人资料推断用户的年龄范围。

年龄推断攻击涉及三个挑战。 首先，需要发现一种可行的方式来收集用户的大量敏感信息。 这些信息被视为必要的地面实况数据。 其次，需要探索微博数据中所有属性之间的相关性，以辅助建立神经网络。 第三，由于目前在微博上没有建立用于推理攻击的神经网络，所以需要建立所提出框架的最有效配置。

**本文有几个贡献。首先，本文发明了一个整合了更快区域卷积神经网络（R-CNN）的新框架，传统CNN和FCNN（全连接神经网络）。这种整合不仅扩大了社交网络中推理攻击的范围，而且提升了现有模型的性能。其次，论文根据图像修改现有的CNN年龄分类器。所提出的框架在分类器中添加用户的属性以使分类器更强大和更健壮。第三，本文阐述了基于属性的完全连接的年龄分类器的配置。虽然有多篇论文关注类似的分类器，但他们都没有详细解释构建的神经网络的结构。第四，本文展示了如何在现实社交网络中收集敏感用户信息的可行方式。第五，比较了神经网络和不同传统机器学习算法在社交网络推理攻击中的性能。第六，通过利用差分隐私（DP）来抵抗推理攻击，提出了一种有效的对策。**

特别是，为了实现年龄段预测的高精度，提出了一种新的框架。 该框架是一些现有神经网络的整合和修改。 一种主要的整合是在图像上结合更快的R-CNN和CNN年龄分类器。 从R-CNN提取的人脸区域提供CNN年龄分类器的输入，可以显着提高攻击性能。 另一个主要整合是CNN年龄分类器和FCNN年龄分类器的组合。 这种集成拓宽了现有模型的范围，使得所提出的框架能够预测用户的年龄范围，而不管用户是否具有有效的简档图像。 一个主要修改是用户的属性注入CNN年龄分类器。 因此，与仅基于图像的现有分类器不同，利用图像和属性来寻求最佳预测精度。

另一方面，本文还给出了基于属性的完全连接的年龄分类器的构造细节。 在现有的研究中很好地说明了基于图像的更快的R-CNN和CNN年龄分类器。 然而，没有关于如何基于属性构建FCNN年龄分类器的详细信息。 本文的讨论有助于理解FCNN如何用于推理攻击，因为不同的FCNN配置在性能上会有显着差异。

**为了更好地评估社交网络中推理攻击的能力，还应用了几种传统的机器学习算法。 具体而言，我们的比较研究选择了决策树，朴素贝叶斯和k近邻（k-NN）。 这些算法已被证明在不同领域执行推理攻击是有效的。 然而，在社交网络领域，很少有研究将性能与神经网络进行比较。**

**最后但同样重要的是，本文还提出了一种有效的防御机制，它利用DP防止基于学习的推理攻击。 评估结果表明，我们的防御机制不仅可以通过降低成功率有效地阻止本文提到的推理攻击，还可以通过寻找相对较大的隐私预算和最少的属性来保护数据的可用性。**

**本文的结构如下。 第2节介绍了最近在社交网络中推理攻击的相关工作。 第3部分详细介绍了提出的框架，解释了该框架如何应用于目标社交网络，并推导出DP防御机制。 第4部分评估攻击框架和DP防御机制的性能。 最后，第5部分总结全文。**

# 相关工作

许多研究[1]，[2]，[3]，[4]，[5]，[6]，[7]，[8]都是基于传统的机器学习算法推断隐藏用户信息。 这些研究一般分为两类，第一类通过部署无监督的机器学习算法将用户分为不同的类别[1，2，3，4]，第二类使用传统的机器学习算法与自然语言处理结合[5]，[6]，[7]，[8]。

**但是，这些研究有两个缺点。 首先，社交网络上的传统机器学习算法的性能普遍较差。 由于社交网络的复杂性，通常有10个以上的公共属性，追求公共属性与目标隐藏属性之间的线性关系通常具有挑战性。 其次，一些研究使用自然语言处理来提高表现。 相反，由于传统的机器学习算法的限制，所添加的技术是耗时的并且本质上也无法提高性能。**

**最近，一些研究[9,10,11,12]已经实现，以利用基于FCNN的推理攻击在社交网络中。 选择FCNN是因为它擅长寻找输入属性和输出属性之间的复杂关系。 当难以用线性公式表达关系时，这一优势变得尤为重要。 但是，现有的研究有两个主要缺点。 首先，这些研究既没有清楚地显示FCNNs的配置，也没有显示研究数据集的详细特征。 由于不同的配置会显着影响社交网络中推理攻击的性能，因此演示FCNN如何构建至关重要。 其次，没有一项研究比较FCNNs和传统机器学习算法的性能。 由于FCNN基本上是一种机器学习算法，因此从全局角度来看，部署这种比较非常重要。**

**最近，CNN一直在迅速发展。 由于CNN的改进，计算机视觉和模式识别有几个突破。 基本上，用于图像处理的CNN包含两个组件：卷积（conv）层和完全连接的层。 Conv层是CNN与其他种类的神经网络区别所在。 这些图层的目的是提取图像的特征。 较低的conv层的任务是提取图像中的曲线或三角形等特征，而较高层的特征是提取更复杂的特征，如面部，汽车或花朵。 在卷积层之后，完全连接的层接管。 像FCNNs中的所有层一样，这些层的目的是进行分类。 在这种情况下，完全连接的图层会将从conv图层提取的要素分类到不同的类别中**

Levi和Hassner [13]提供了使用CNN根据社交网络中的个人资料图像对用户的年龄和性别进行分类的细节。 它表明，通过使用CNN学习代表性，可以在自动年龄和性别分类上获得显着的性能提升。 但是，分类器仅在用户拥有包含清晰面部的配置文件图像时才起作用。 这并非总是如此，因为许多用户在上传个人资料照片时存在隐私担忧。

**Ren等人 [14]引入了更快的R-CNN。 原R-CNN的目的是解决目标检测问题。 给定一个图像，可以在所有对象上绘制边界框。 该过程可以分为两个通用组件，区域定位步骤和分类步骤。 但是，最初的R-CNN包含复杂的管道，运行速度相对较慢。 速度更快的RCNN能够解决复杂性问题，使模型更有效地运行。 目前，更快的R-CNN已经成为图像中物体检测的金标准。 因此，本文提出的框架的第一步使用更快的R-CNN。**

DP是一种流行的隐私保护机制，广泛应用于各个领域的研究。 最受欢迎的领域之一是推荐系统。 McSherry和Mironov [15]探讨了将DP应用于Netflix Prize竞赛中领先方法的方法，而不会显着降低准确性。 DP也逐渐应用于社交网络。 高等人[16] 将DP应用于社交网络中的分层随机图结构以保证更强的去匿名保护。 上述两项工作均着重于大规模用户身份匿名化保护，并且可能不适用于小型数据集。 相比之下，我们的方法可以保护可伸缩社交网络数据集的敏感属性免受基于学习的推理攻击。

# 方法

## 3.1 系统总览

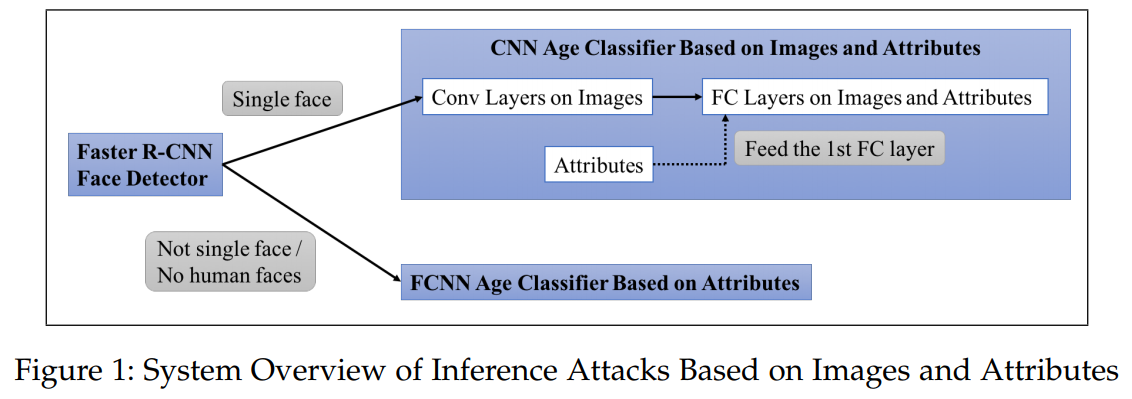
图1显示了该系统的概况。它包含三个主要部分，即更快的R-CNN人脸检测器，基于图像和属性的CNN年龄分类器以及基于属性的FCNN年龄分类器。更快的R-CNN检测用户简档图像中人脸的存在。如果R-CNN在图像中检测到合法的单个人脸，则使用基于图像和属性的CNN年龄分类器。如果R-CNN无法检测到任何人脸或图像中有多个人脸，则使用基于属性的FCNN年龄分类器。在基于图像和属性的CNN年龄分类器中，有两个基本组件：conv层和完全连接的层。从更快的RCNN提取的单幅图像是conv层的输入。这些conv图层仅适用于脸部图像。 conv图层的输出是一个包含年龄分类关键特征的向量。然后，将性别，工作信息和教育信息等来自同一用户个人资料的属性与来自conv层的输出相结合，成为完全连接的图层的输入。之后，完全连接的图层将同时处理图像特征和属性以分类用户的年龄。

由于提出的框架是对现有模型的整合和修改，所以模型如何整合以及修改背后的动机如3.2节至3.4节所示。大规模的地面实况数据必须被抓取以训练神经网络，因此第3.5节展示了如何从真实社交网络收集数据。 3.6节打算寻求用户公共属性和目标隐藏属性之间的关系。在这一步中，涉及相关矩阵来理解可用数据的特征。由于现有研究没有显示FCNN针对社交网络攻击的详细结构，因此第3.7节通过阐述一个基于公共用户信息的推理攻击的可行且有效的FCNN来解决这一突出问题。 3.8节展示了一些典型的传统机器学习算法的实现。这些算法被用来比较神经网络的性能。具体来说，使用了决策树，朴素贝叶斯算法和k-近邻算法。这些算法被证明对于信息攻击是有效的。 3.9节描述了本文中使用的交叉验证。交叉验证是分析机器学习算法的正确分类百分比的公正方法之一。最后，第3.10节提供了DP的防御机制

## 更快的R-CNN人脸检测器和CNN年龄分级器的集成方法

基于图像的更快的R-CNN人脸检测器和CNN年龄分类器是成熟的模型。 更快速的R-CNN人脸检测器可以获得理想的结果，从而准确有效地提取图像中的人脸。 如果包含单张脸的图像是输入，则基于图像的CNN年龄分类器也可以实现很好的结果。 本文的主要贡献之一是两种模型都被集成来执行推理攻击。

部署这种集成有两个好处。 首先，R-CNN人脸检测器可以滤除不包含人脸的轮廓图像。 因此，可节省时间以跳过图像上的复杂CNN分类器，代之以执行轻量级FCNN分类器。 其次，整合可以提高CNN年龄分类器在图像上的准确性。 由于原始分类器仅在人脸图像上进行训练，因此当输入清晰的人脸图像时效果最佳。 更快的R-CNN人脸检测器可确保每个送入CNN分类器的图像都包含清晰的单个人脸。 原因是R-CNN人脸检测器能够提取图像中的所有人脸区域，同时消除背景噪声和其他物体的噪声。 然后，如果只有一个合法人脸，则提取的人脸区域而不是原始图像成为CNN分类器的一个输入。



## CNN年龄分类器与用户属性的整合

原始的CNN年龄分类器仅基于图像。 虽然分类器本身已经提供了准确的预测，但是对于这个成熟的模型可以进行改进。 本文的另一个主要贡献是用户的属性融入了传统的CNN模型。 因此，通过将图像和属性作为输入，CNN年龄分类器变得更加强大和健壮。

为了将属性注入到现有的分类器中，使用了迁移学习。 迁移学习是一个采用预先训练的模型并调整模型以针对类似问题的过程。 如前所述，CNN模型包含conv层和完全连接的层。 在这种情况下，现有CNN中的conv层中的所有权重都被保存。 然后，将通过conv图层将相同用户属性和从配置文件图像中提取的要素进行连接。 串联的矢量成为完全连接的层的输入。 训练时，只需确定完全连接层中的权重。

采取这种方法有两个好处。 首先，如果从图像预测的信息不足，来自用户的属性可以补充他的脸部图像。 其次，即使用户的图像已经提供了足够的信息，用户的属性仍然具有提高准确性的巨大潜力。 上述两个优点背后的原因是，属性为用户带来了宝贵的附加信息。 不仅仅依赖于一个源文件图像，两个来源增加了更好地预测用户隐藏信息的机会。

## CNN年龄分级器和FCNN年龄分级器的集成方法

用户的个人资料图像可以提供丰富的隐藏属性信息。 然而，用户不会将合法的个人资料图像上传到社交网络上是相当普遍的。 这种现象背后的原因之一是用户试图保护他们的隐私。 另一个原因是大多数社交网络不验证用户的个人资料图像。 为使目标系统成为更通用的场景，本文的另一个主要贡献是仅基于属性的FCNN分类器和基于CNN分类器的图像和属性的融合。

这种集成有两个好处。首先，系统可以适应更一般的情况，因为不是每个用户都有一个配置文件映像。尽管用户可以上传图像，但该图像可能与用户的脸部无关。请注意，用户选择组图像作为配置文件图像并不罕见。在这种情况下，这些图像仍然被视为无效，因为在训练阶段很难验证组图像中的哪个人脸是用户。其次，整合可以提高预测的准确性。一方面，如果仅考虑基于图像和属性的CNN分类器，图像可能包含妨碍准确预测的不利信息。例如，如果用户的个人资料图像只是一朵花，使用此照片来训练CNN分类器会起到相反的作用。另一方面，如果仅考虑基于属性的FCNN分类器，浪费有价值的图像，并且预测准确性可能受到负面影响。因此，CNN和FCNN分类器的整合对于提高预测精度具有重大的潜力。

## 数据采集

选择新浪微博作为收集用户数据的主要网站有两个原因。 首先，它是中国最大的在线社交网络。 截至2015年底，用户数量超过2.22亿，日用户数量达到1亿。 每天发布大约1亿条消息。 其次，微博的数据质量非常高。 由于中国政府的政策，微博用户必须注册其真实身份。 具体而言，网站在注册过程中要求提供有效的手机号码，中国政府强制要求所有手机号码必须使用经过验证身份的实名。 换句话说，微博中的用户数据质量是可以承诺的。 百度百科是一个中文，合作，基于网络的百科全书。 它包含了中国所有公众，知名和有影响力的人的准确个人信息。

由于推理攻击是通过神经网络进行的，所以要训练的数据是必不可少的。 大量的地面实况数据必须从互联网上检索。 但是，包含私人个人信息的数据很难收集。 为了解决这个问题，微博和百度百科的数据都被抓取。 然后，连接从百度百科提取的年龄信息和从微博获取的个人资料信息。 两个网站的组合构成足够的包含敏感年龄信息的地面实况数据。 因此，收集了来自2030人的个人信息。

个人资料图片仅通过百度百科收集。 原因在于，为了确保权威性，百度百科保证每个人都有合法清晰的个人形象。 否则，百度百科不会为该人提供任何个人资料图片。 因此，所有收集到的图像都可以视为地面真实图像，并能正确反映人们的年龄

除配置文件图像外，对于CNN和FCNN中的属性部分，还考虑了以下属性：配置文件位置，性别，博客的存在，口号的长度，注册年份，工作信息的存在，教育信息的存在， 以下和帖子数量。 请注意，为了保持问题的普遍性，不考虑追随者的数量，因为名人通常比普通用户拥有更多的追随者。 针对不同的年龄段设置18个级别，年龄从5-10岁到90-95岁不等。

## 相关矩阵

相关矩阵可以表示所选属性之间的关系。 它显示了任何两个属性之间的依赖关系。 方程（1）显示了两个变量x和y之间的相关矩阵。 x是自变量。 y是因变量。 n是样本中数据点的数量。 和分别是x和y的均值。 Sx和Sy分别是x和y的样本标准差。 相关值r（x，y）是-1到1之间的数字。在正相关中，随着x增加，y增加。 在负相关中，随着x增加，y减小。 如果该值接近0，则意味着x和y是松散关联的。

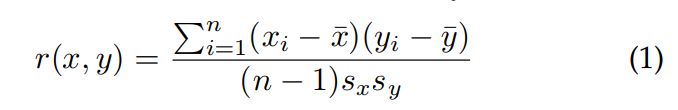
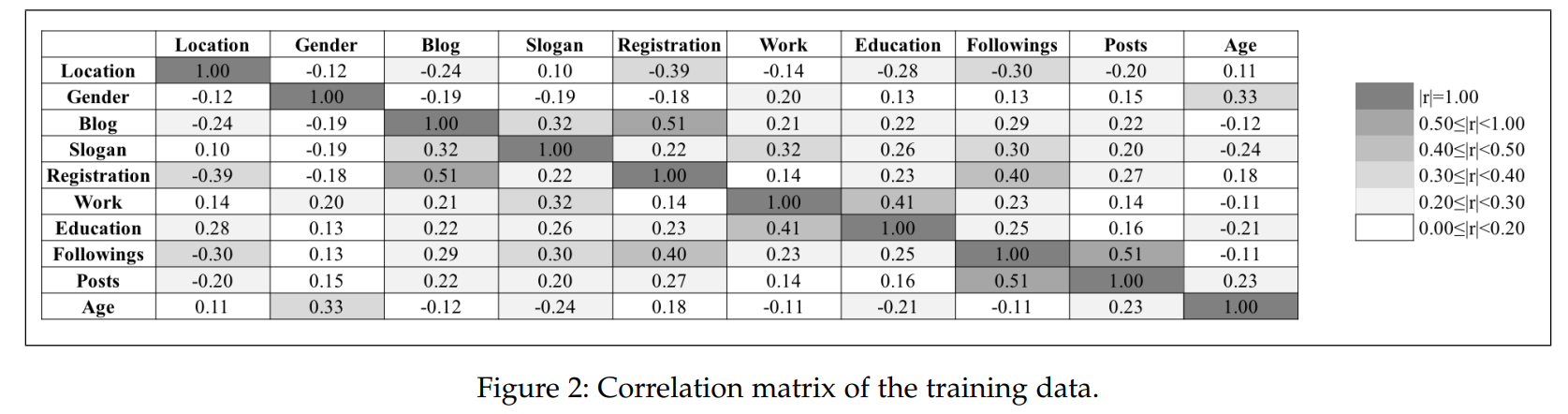


图2是相关矩阵。 如果两个属性具有较高的相关性，则该单元格的背景颜色较暗。 在图2中，所有公共属性之间的关系都很弱，所有的公共属性与私有属性的关系都很弱。 由于所有的属性并不相互独立，因此它是实现神经网络的理想环境。 原因是所有的公共属性都可以有效地帮助对方学习与目标私人属性的隐藏关系。

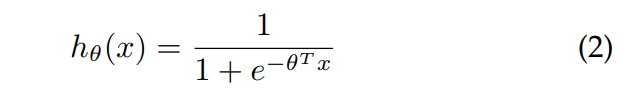


## 全连接神经网络结构

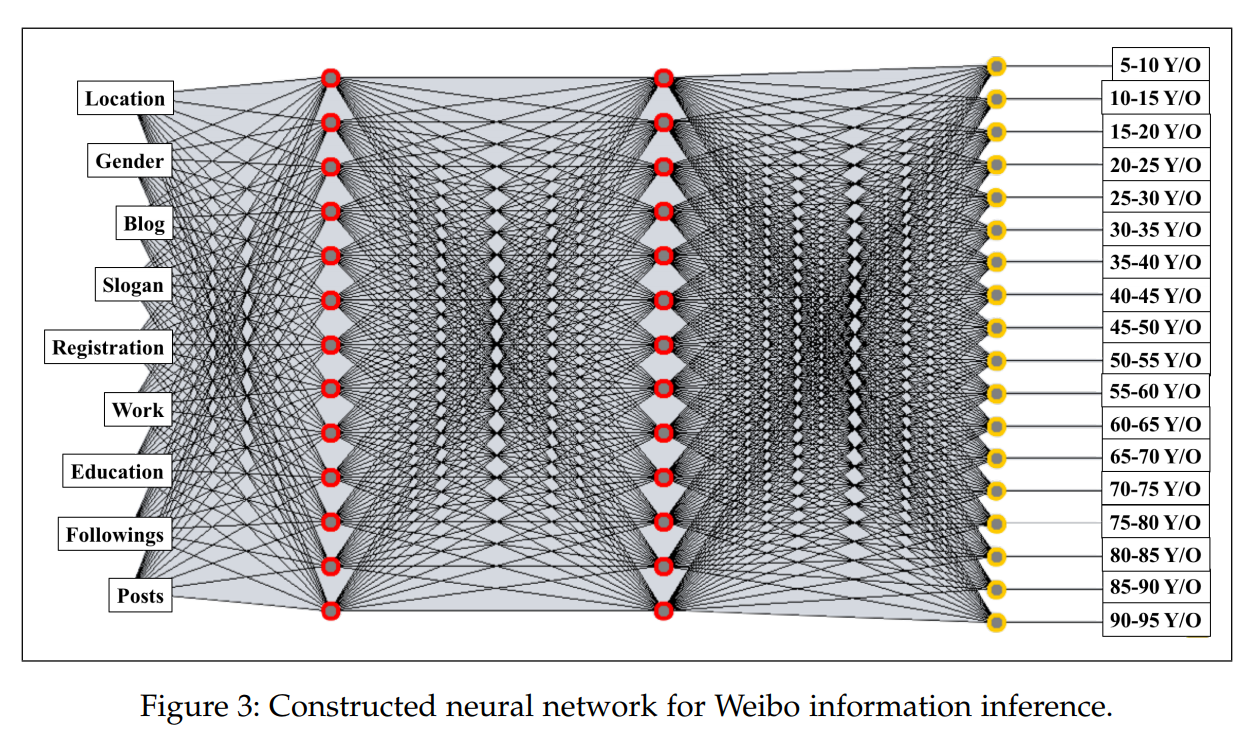
基于图像的R-CNN人脸检测器和CNN年龄分类器在[13]，[14]中清楚地展现出来。 然而，现有的论文没有清楚地显示FCNN年龄分类器在属性上的配置。 不同的配置会显着影响FCNN的性能，因此显示FCNN建设的细节是迫切需要的。

**选择FCNN有两个原因。 首先，根据问题的性质，属性的数据类型很简单，只有少数已知属性和几个目标类。 FCNN善于处理这种设置。 其次，FCNN计算效率高，易于实现。 因此，调整网络参数以寻求问题的最佳配置是灵活的**

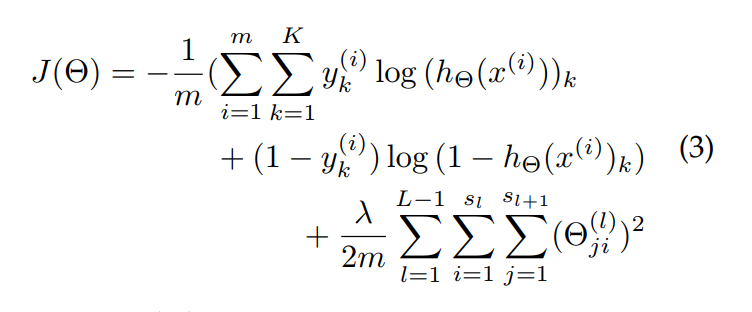
在典型的FCNN中，有几层，每层都有几个神经元。 神经元本质上是一个逻辑单元，由逻辑激活函数表示。 在本文中，使用（2）中所示的S形激活函数。 在（2）中，hθ（x）是需要表达的目标假设。 θ和x分别是不确定的权重向量和已知的属性向量。 对于输入图层，由于测量了用户的九个公共属性，因此有九个单位。 对于输出图层，有18个单元，因为对于不同的年龄段设置了18个类。



对于FCNNs，关键是要找到每层最佳层数和单元数，以适当地解决问题。 由于没有强制性规则来定义这些值，所以最好的方法是通过反复试验和问题的特征。 经过多次尝试，神经网络的最佳结构是有2个隐藏层，每层有13个单元。 神经网络的完整结构如图3所示。



（3）中显示了标准的从逻辑回归中推导出来神经网络代价函数，。（3）中的之前的部分是计算训练数据与当前假设结果之间的误差。 以后的部分是防止过拟合的正则化部分。 具体而言，m是训练数据的数量。 K是输出类的数量。 （x（i），y（i））表示训练数据集中的条目。 λ是控制过拟合的正则化参数。 L是网络中的层数。 Sl是第l层中的单位数，不包括偏差单位。

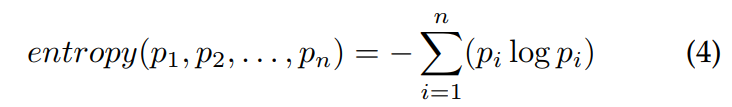


为了最小化J（Θ），实施反向传播算法来计算。 注意到要实现最佳性能，还要执行梯度检查和随机初始化

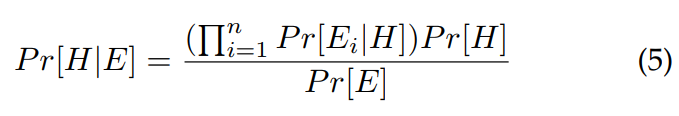
## 传统机器学习方法实现

在本节中，我们检查了几种传统的机器学习算法，并将其与所提议的框架的结果进行比较。 选择这些算法是因为它们已被广泛证明对推理攻击有效。 特别是，实现了决策树，朴素贝叶斯和k-近邻算法。

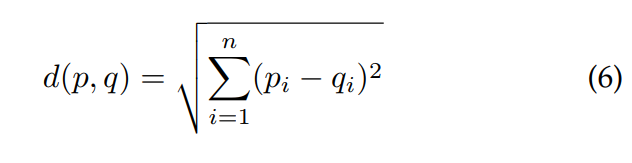
第一个应用的机器学习算法是决策树。 决策树采用自顶向下和分而治之的策略。 关键是要知道哪个是最好的属性，目的是获得最小的树。 从技术上讲，所选属性应该具有最大的信息增益。 信息增益计算为拆分前的分布熵与拆分后的分布熵之间的差值。 在这里，信息是以比特为单位的。 （4）中定义熵，p表示概率。



第二种机器学习算法是N朴素贝叶斯。 该算法假定知道一个属性的值并没有说明另一个属性的值。 方程（5）显示了贝叶斯定理。 P r [H | E]是给定证据E的事件H的概率，也被称为H的后验概率。P r [H]是H的先验概率



第三种算法是k近邻。 经过多次尝试，k = 20最适合这个问题。 一个对象通过其邻居的多数投票进行分类，该对象被分配给它的k个最近邻居中最常见的类。 使用相似性函数来搜索训练集。 本文实现了欧氏距离。 距离d（p，q）如（6）所示，其中p和q是两个n维数据点



## 交叉验证

交叉验证是评估统计分析结果如何推广到独立数据集的模型验证技术。 它主要用于目标是预测的环境中，并且人们希望估计预测模型在实践中的准确度。 在预测问题中，模型包含运行训练的已知数据的数据集，以及模型测试的未知数据的数据集。 交叉验证的目标是限制过度拟合问题，从而深入了解模型如何推广到独立数据集

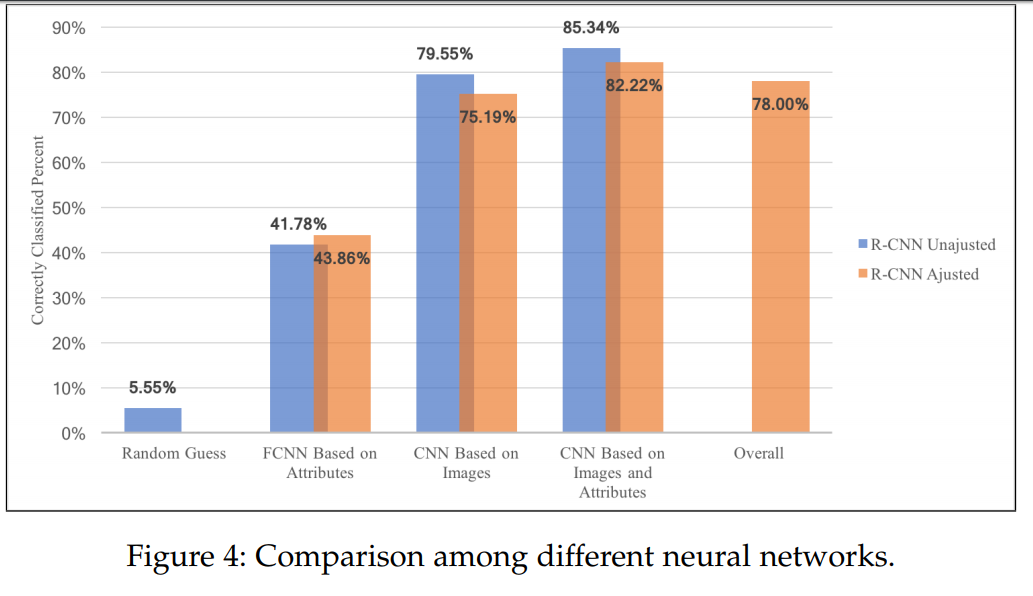
在本文中，交叉验证设置为10个包。 这意味着数据分成10部分，每部分有203个独立项。 之后，使用9个部分进行训练，并且使用1个部分进行测试。 这个过程通过选择不同的部分来重复10次来训练和测试，以使结果真实地反映事实。

## 防御差异隐私

# 结果和讨论

## 4.1 综合框架评估

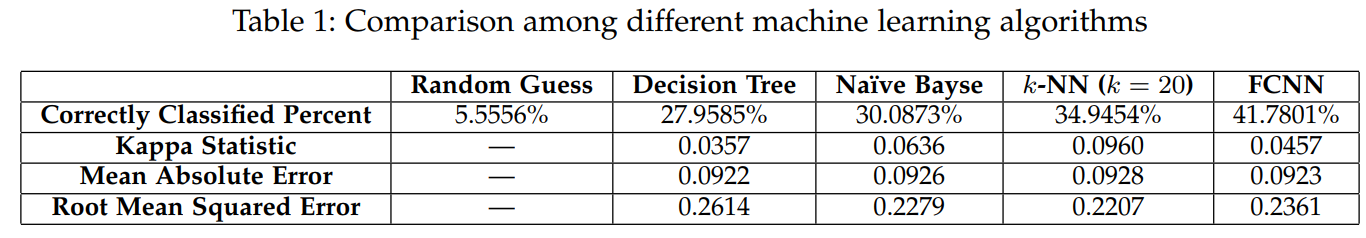
图4显示了提出的框架不同组成部分之间的评估和比较。 可以清楚地看到，基于属性的FCNN具有很强的预测用户在社交网络中的年龄范围的能力。 结果大约是随机猜测结果的8倍。 因此，证明了公共属性通过单独使用FCNN来预测用户的私有属性具有很大的潜力。 然而，由于公共属性中包含的信息仍然有限，预测率仍低于50％。



来自CNN的基于图4中图像的结果显示了仅利用用户的简档图像的推理攻击的结果。 这个结果与[13]的结果一致。 但是，在本文中，现有模型已被修改。 属性因素通过传递学习技术融合在CNN中。 这项技术的过程在3.3节详细说明。 因此，从图4可以看出合理的提升。对于R-CNN未调整的情况，预测率从79.55％上升到85.34％，增加了7.28％。 性能提升背后有两个原因。 首先，如果图像缺乏预测用户年龄范围的能力，则添加的属性信息可以补偿图像以提供对用户更好的洞察。 其次，如果图像包含关于用户年龄的足够信息，则添加的属性有能力加强信息并帮助推理攻击。

图4还显示了R-CNN调整前和R-CNN调整前的结果之间的比较。 R-CNN的引入是为了增加所提议框架的灵活性。 因此，无论数据集中具有有效概要文件图像的用户的百分比如何，框架都可以调整数据集并最大化预测准确性。 为了评估灵活性，测量R-CNN调整前和R-CNN调整后的绝对差值。 如果差异很小，则意味着不仅框架灵活处理不同的数据集，而且总体预测率也可以得到保证。 在图4中可以看出，在所有情况下，波动并不大，说明框架的预测能力是有前途的和强大的。 可以看出，框架总体预测率高达78.00％。

表1显示了构建的FCNN以及其他几种机器学习算法的结果。 有四个指标正在考虑，这是正确的分类百分比，卡帕统计，平均绝对误差和均方根误差。 对于正确分类的百分比，实施交叉验证以实现令人信服的结果。 Kappa统计量是衡量定性项目的评价人间协议。 平均绝对误差测量一组预测中误差的平均值，而不考虑其方向。 均方根误差是一个二次评分规则，也可以衡量误差的平均值。



正确分类的百分比是反映每种算法性能的最重要的指标。 来自每个传统机器学习算法的百分比可以达到大约6倍于随机猜测的百分比。 由于数据是直接从Internet上抓取的，因此数据的分布不规范。 当k设置为相对较大的数字时，k-NN具有处理噪声数据的最佳能力。 因此，k-NN在三种传统机器学习算法中表现最好。 此外，从问题的本质来看，很难找到敏感年龄信息与所有不敏感属性组合之间的关系。 图2证实了相同的难度，因为所有可用的属性具有非常有限的相关性。 从这个角度来看，三种机器学习算法提供了相对合理和有效的结果。

表1还表明，FCNN约为随机猜测精度的8倍。当将FCNN的结果与任何传统的机器学习算法的结果进行比较时，显然FCNN优于所有三种传统算法。有关FCNN为什么表现更好的原因有三个。首先，FCNN善于在已知属性和未知属性之间寻找复杂而微妙的关系。在社交网络推理攻击中，已知属性非常松散地连接到目标隐藏属性。其次，FCNN的假设是复杂的，不限于线性表达式。由于传统的机器学习算法只能追求由线性表达式表达的关系，FCNN打破了这种限制，并能够寻求非线性关系。第三，由于社交网络数据是嘈杂的，FCNN比其他任何传统的机器学习算法具有更强的处理噪声的能力。由于这三个原因，构建的FCNN具有更高的正确分类百分比来推断社交网络的隐藏属性

从整个图4和表1的结果看，当前的社交网络存在相当大的隐私问题。 如今，社交网络提供商尽力保护用户的私人信息，并为用户提供控制其个人资料的权利。 然而，从本文的结果来看，用户的私人信息仍然是脆弱的。 一方面，像FCNN攻击这样简单的神经网络攻击可以实现将近50％的预测精度。 另一方面，像本文提出的框架这样复杂的神经网络攻击可以达到几乎80％的整体预测精度。 请注意，随着现代计算的发展，复杂的神经网络不需要难以承受的时间来执行。 这种情况进一步使用户的隐私处于风险之中。 总之，当前的社交网络隐私仍然容易受到现代神经网络引发的推理攻击的影响。

# 总结

本文提出了一个新的社交网络推理攻击框架。所提出的框架巧妙地整合并修改了现有的最先进的CNN模型。因此，无论用户是否拥有合法的个人资料图片，它都可以适应更广泛的推理攻击场景。实验显示总体预测准确率为78.00％，几乎是随机猜测准确度的16倍。实验还表明，框架可以将推断攻击的现有高精度CNN提高7.28％。除了新框架外，本文还提供了FCNN推理攻击的详细配置，这在现有研究中通常是缺失的。此外，传统的机器学习算法被用来比较来自构建的FCNN的结果。最后但并非最不重要的是，DP被用作防御机制。总的来说，本文令人信服地表明，当前的社交网络容易受到现代神经网络引发的信息推理攻击

本文有两个主要限制。 首先，只有一个社交网络经过培训和测试。 其次，只考虑一个有针对性的私人属性，即年龄范围。 在未来的工作中，将探索更多的社交网络和更具针对性的属性，以进一步验证所提议的框架的性能，并在社交网络中展示更多不同的推理攻击。