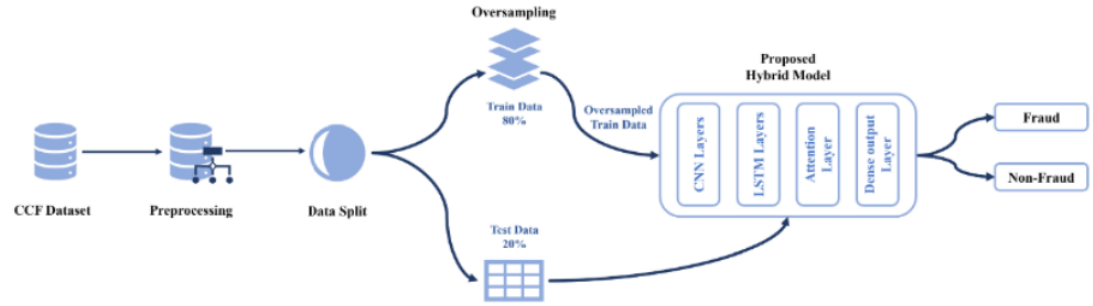
使用数据集：

Credit_Card_Fraud_Detection

该数据集共有 284807 条交易记录，其中有 492 条为诈骗交易，数据分布高度不平衡，诈骗记录占所有交易的 0.172%。

每条记录包含时间，金额，功能（V1-V28）和类，前三者是特征，其中时间是每个事务与数据集中第一个事务之间经过的秒数。最后的类是标签，在欺诈交易的情况下取 1，正常交易取 0。

论文神经网络训练流程：



可见就是一个二分类问题。数据预处理后，切分为训练集和测试集（应该还会从训练集切分出验证集），神经网络包含四个主要部分：卷积层，LSTM 层，注意力层和输出层。最后的输出是二分类。

数据集的处理：

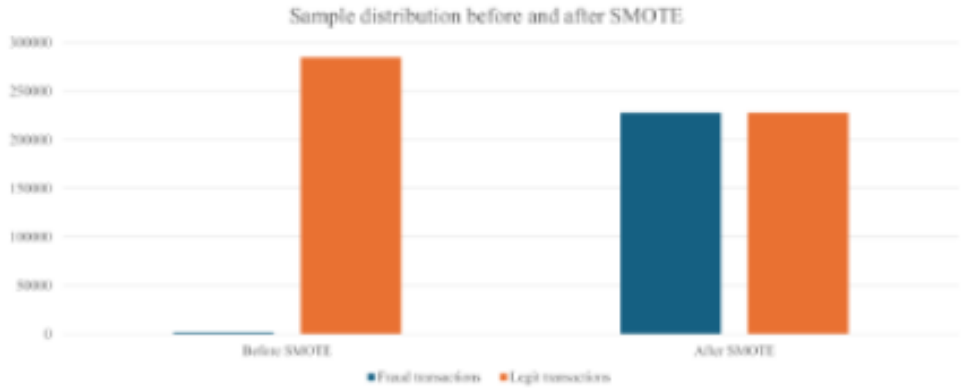
首先是数据的归一化。通常而言会使用减去平均值，除以标准差，将数据压缩到标准正态分布上，本篇论文使用了：

$$X_{\text{norm}} = \frac{X - X_{\text{min}}}{X_{\text{max}} - X_{\text{min}}}$$

这是另外一种缩放方式，可以将数据压缩到[0,1]，但是对于稀疏数据可能会破坏稀疏性。

处理类不平衡：SMOTE

在原先的数据集基础上使用 SMOTE 生成数据集，实现数据集的平衡，它通过在样本中划一条线，在沿线的某一点绘制新样本来合成数据。处理后样本分布：



具体参照：SMOTE: Synthetic Minority Over-sampling Technique

神经网络层：

由卷积神经网络层，LSTM 层和注意力层组合。卷积层使用 3*3 的窗口和 relu 激活函数，并使用 2*2，步幅为 2 的最大池化层缩小特征图。

LSTM 层：

处理长期依赖关系（时间特征），包含输入门，忘记门和输出门，在长序列中保留相关信息。

注意力层:

增强模型专注于输入序列重要部分的能力。