Solution Architecture- Web Phishing Detection
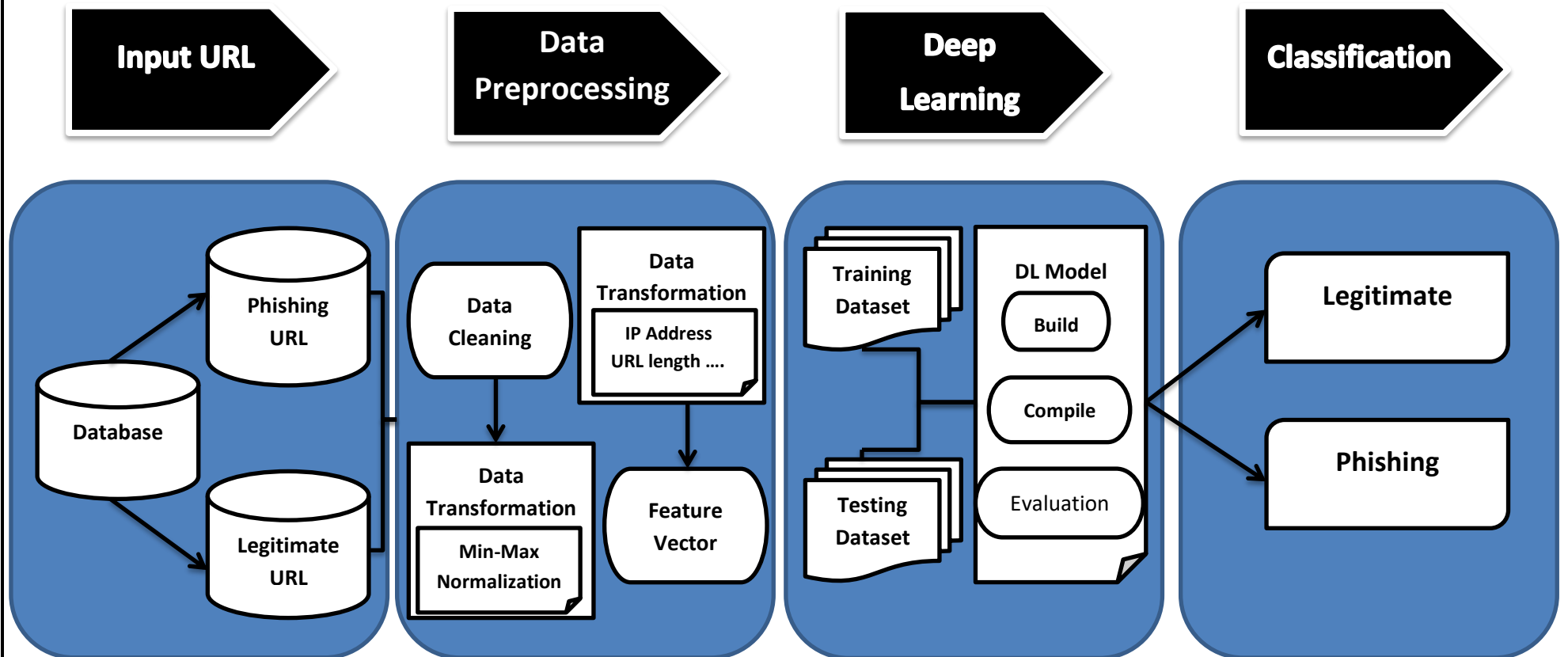
Team ID: PNT2022TMID32768

Team Members:

Team Leader:　　　T.Mohamed Musaraf

Team Member 1:　　B.Prasanna Venkatesh

Team Member 2:　　N.RamKumar

Team Member 3:　　S.Srinivasan

ARCHITECTURE

Input URL

Data Preprocessing

Deep Learning

Classification

Database

Phishing URL

Legitimate URL

Data Cleaning

Data Transformation
IP Address URL length ....

Data Transformation
Min-Max Normalization

Feature Vector

Training Dataset

Testing Dataset

DL Model
Build
Compile
Evaluation

Legitimate

Phishing

FLOW CHART:

Webpage Feature Generation

Construct DOM tree of Webpage

Get URL of Webpage

Extracted Features

Hyperlink Features

URL Features

Textual content Features

Feature Vectorization

Training Set

Testing Set

Machine Learning

Detection Module

Decision

Legitimate

Phishing

Preprocessing Phase

Detection Phase