



中南大學  
CENTRAL SOUTH UNIVERSITY

# XXXX 论文

题目	XXXXXXXX
学生姓名	XX
学 号	XXXXXXXXXX
专业班级	XXXXXX
指导教师	XXXX
学 院	XXXXXX
完成时间	XXXXXXXX



中南大学  
CENTRAL SOUTH UNIVERSITY

# 论文题目

学生姓名：\_\_\_\_\_XXX\_\_\_\_\_

学    院：\_\_\_\_\_XXX 学院\_\_\_\_\_

专业班级：\_\_\_\_\_XXXX\_\_\_\_\_

指导老师：\_\_\_\_\_XXX 教授\_\_\_\_\_

二〇一九年十二月

## 摘要

提供了三个封面，按需

正文 宋体 小四，20 磅行距 最小值

标题从三号字体依次递减 最多三级标题

图片下方标题由文本框加入，设置布局格式为上下型环绕

## 目录

摘要 .....	2
目录 .....	3
第一章 绪论 .....	4
1.1 背景 .....	4
1.2 研究现状 .....	4
第二章 xxxx 简介 .....	4
2.1 XXXX .....	4
2.2 XXXX .....	4
第三章 xxxxx .....	5
3.1 xxxxxx .....	5
3.2 xxxx .....	5
3.2.1 xxx .....	5
3.2.2 xxx .....	6
第四章 结束语 .....	6
参考文献 .....	7

## 第一章 绪论

### 1.1 背景

XXXXXXX。

### 1.2 研究现状

## 第二章 XXXX 简介

### 2.1 XXXX

XXXXX。

### 2.2 XXXX

XXXX[l]。

$$\arg \max_{\mathcal{D}_p} \mathcal{W}(\mathcal{D}', \theta_p^*)$$

## 第三章 XXXXX

### 3.1 XXXXXX

---

#### Algorithm 1 Poisoning Attack Algorithm

---

**Input:**  $\mathcal{D} = \mathcal{D}_{\text{tr}}$  (white-box) or  $\mathcal{D}'_{\text{tr}}$  (black-box),  $\mathcal{D}'$ ,  $\mathcal{L}$ ,  $\mathcal{W}$ , the initial poisoning attack samples  $\mathcal{D}_p^{(0)} = (\mathbf{x}_c, y_c)_{c=1}^p$ , a small positive constant  $\varepsilon$ .

```

1:  $i \leftarrow 0$  (iteration counter)
2:  $\boldsymbol{\theta}^{(i)} \leftarrow \arg \min_{\boldsymbol{\theta}} \mathcal{L}(\mathcal{D} \cup \mathcal{D}_p^{(i)}, \boldsymbol{\theta})$ 
3: repeat
4:    $w^{(i)} \leftarrow \mathcal{W}(\mathcal{D}', \boldsymbol{\theta}^{(i)})$ 
5:    $\boldsymbol{\theta}^{(i+1)} \leftarrow \boldsymbol{\theta}^{(i)}$ 
6:   for  $c = 1, \dots, p$  do
7:      $\mathbf{x}_c^{(i+1)} \leftarrow \text{line\_search} \left( \mathbf{x}_c^{(i)}, \nabla_{\mathbf{x}_c} \mathcal{W}(\mathcal{D}', \boldsymbol{\theta}^{(i+1)}) \right)$ 
8:      $\boldsymbol{\theta}^{(i+1)} \leftarrow \arg \min_{\boldsymbol{\theta}} \mathcal{L}(\mathcal{D} \cup \mathcal{D}_p^{(i+1)}, \boldsymbol{\theta})$ 
9:      $w^{(i+1)} \leftarrow \mathcal{W}(\mathcal{D}', \boldsymbol{\theta}^{(i+1)})$ 
10:   $i \leftarrow i + 1$ 
11: until  $|w^{(i)} - w^{(i-1)}| < \varepsilon$ 

```

**Output:** the final poisoning attack samples  $\mathcal{D}_p \leftarrow \mathcal{D}_p^{(i)}$

---

图一、XXXX 算法

### 3.2 XXXX

#### 3.2.1 XXX

XXXXXX。

### 3.2.2 xxx

Xxxx

Xxxx

## 第四章 结束语

Xxxxxxx。

感谢 xxxxx。

## 参考文献

- [1] S. Alfeld, X. Zhu, and P. Barford. Data poisoning attacks against autoregressive models. In AACL, 2016.