



Prosecutorial Outcome Prediction with LoRA and QLoRA

**KUO-CHUN CHIEN, CHIA-HUI CHANG, HUAI-HSUAN
HUANG, JO-CHI KUNG**

Web Intelligence and Data Mining (WIDM) Lab
Department of Computer Science and Information Engineering
National Central University, Taiwan



Research Background

1 Traditional LJP Focus

Traditional Legal Judgment Prediction (LJP) focuses on court judgments

3 Problem

Current LJP research overlooks the crucial prosecutor's indictment

2 Key Innovation

Shifting focus to prosecutor's needs

4 Goal

Improve efficiency and fairness in legal decision-making

Core Prediction Tasks

Four innovative tasks designed for prosecutors:

1. Reasons for non-punishment

Predicting circumstances where punishment may not be appropriate

2. Imprisonment prediction

Predict whether the court's verdict includes a prison term

3. Fine prediction

Predict whether the court's verdict includes Fine

4. Penalty type determination

Predict the type of punishment, whether it includes prison time, fines, or both



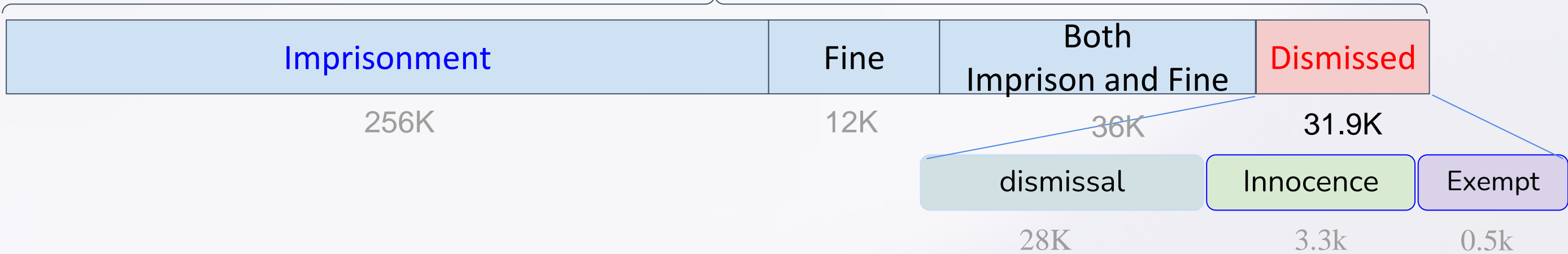
Dataset: TWPJD

Source	Taiwan's Ministry of Justice
Content	Criminal cases with indictments and first-instance judgments
Total Size	336,773 cases
Training	235,740 (70%)
Validation	33,678 (10%)
Testing	67,355 (20%)
Average Length	367.37



Mapping Between Indictments and Verdicts

- 336K Criminal cases



- Task Flow



Technical Approach: Multi-Task Learning (MTL) Framework

1 Simultaneous Learning

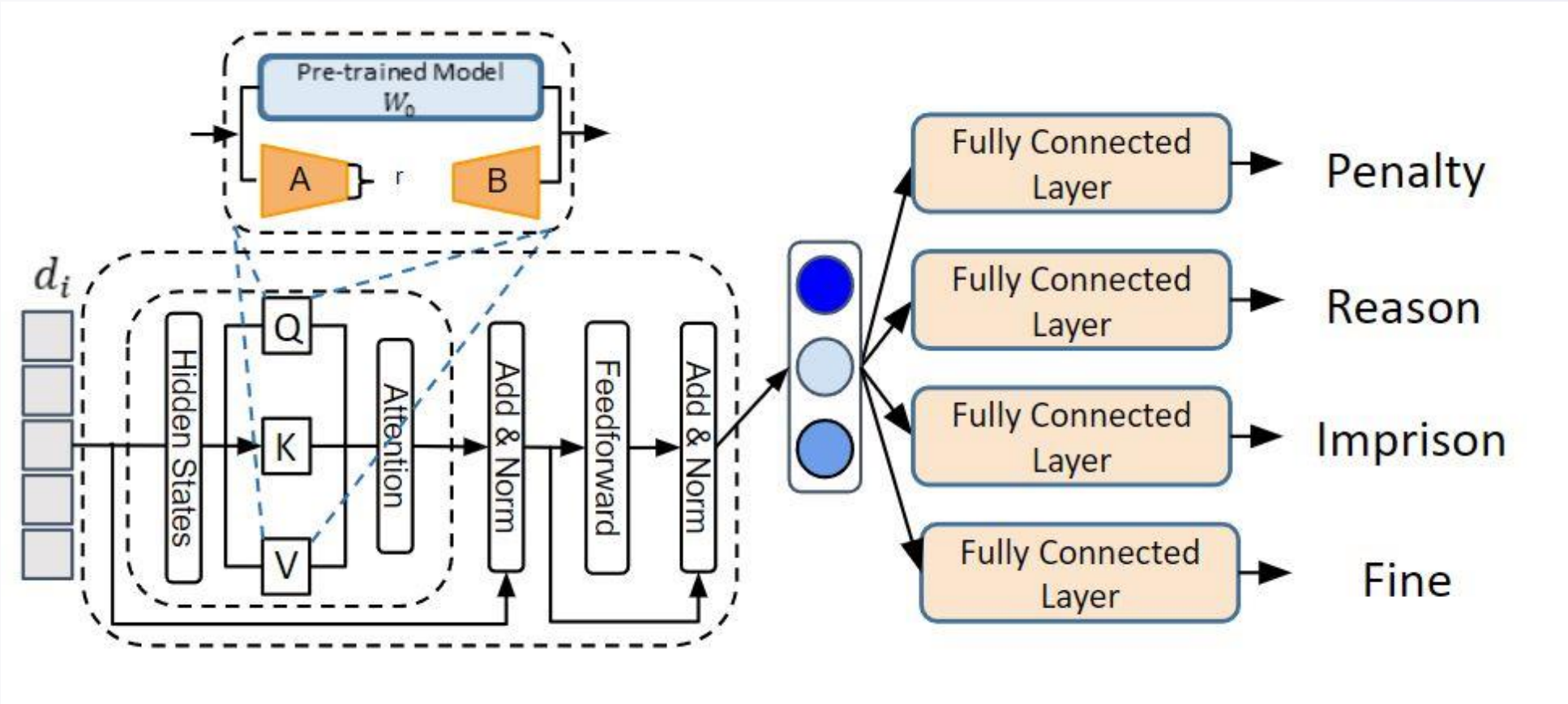
Learns multiple related tasks simultaneously

2 Task Associations

Leverages task associations

3 Enhanced Techniques

Low-Rank Adaptation (LoRA)
Quantized LoRA (QLoRA)



The Performance of TWPJD Dataset

Sub-task	Reason		Imprison		Fine		Penalty	
Model/Metric	Mi-F	Ma-F	Mi-F	Ma-F	Mi-F	Ma-F	Mi-F	Ma-F
Bert_Single	0.909	0.369	0.884	0.735	0.881	0.684	0.793	0.615
Bert	0.912	0.362	0.882	0.729	0.884	0.726	0.804	0.614
Bert+LoRA	0.912	0.372	0.894	0.758	0.886	0.714	0.802	0.613
Bert+QLoRA	0.908	0.382	0.891	0.777	0.888	0.741	0.802	0.640
Topjudge	0.913	0.376	0.89	0.769	0.882	0.753	0.798	0.645
Topjudge+LoRA	0.915	0.365	0.894	0.765	0.889	0.742	0.809	0.617
Topjudge+QLoRA	0.915	0.371	0.896	0.768	0.891	0.733	0.807	0.616
LLAMA+QLoRA	0.9	0.374	0.889	0.742	0.882	0.705	0.793	0.577

Training information of TWPJD dataset					
Model	Batch size	All Params	Trained Params	Training Time	Used GPU Memory
Bert	8	102M	102M	1hr17mins	6,394MB
Bert+LoRA	8	102M	0.3M	38mins	4,926MB
Bert+QLoRA	8	102M	0.3M	48mins	2,751MB
Topjudge	8	244M	244M	1hr26mins	6,486MB
Topjudge+LoRA	8	244M	141M	42mins	5,063MB
Topjudge+QLoRA	8	244M	141M	53mins	2,864MB
LLAMA+QLoRA	4	6706M	4M	13hrs14mins	11.6G



Challenges in "Reason" Prediction

1

1. Evidence assessment

Difficulty in evaluating the strength and relevance of evidence

3

3. Special circumstances handling

Addressing unique factors that may influence legal decisions

2

2. Legal principle application

Challenges in applying complex legal principles to specific cases

4

4. Complex procedure understanding

Grasping intricate legal procedures and their implications

Error Analysis Examples

Case 1: Evidence Insufficiency

Indictment: False accusation and weapon possession

Actual Outcome: Not guilty

Model Prediction: Guilty

Issue: Model couldn't assess evidence sufficiency

Case 2: Mental Health Factor

Indictment: Illegal weapon possession

Actual Outcome: Exemption from punishment

Model Prediction: Guilty

Issue: Mental illness not mentioned in indictment

Case 3: Procedural Requirements

Indictment: Drug use based on positive test

Actual Outcome: Case dismissed

Model Prediction: Guilty

Issue: Model couldn't recognize procedural requirement for rehabilitation

Future Directions

- 1. Data Enhancement**
 - Focus on "reason" prediction
 - More comprehensive case information
- 2. Model Optimization**
 - Advanced language models
 - Improved fine-tuning techniques



THANK YOU