根据 NHTSA 数据训练的事故预测模型

SQ

.1 历史模型

ORIM 历史模型是从人类事发交通事故力总结出来的隐式风险规则。

1.1 事故模型生成

过程见图 47。

图 47

1.2 训练数据来源

包括（不限于）：美国 NHTSA 数据库、德国 GIDAS 数据库、日本 ITARDA 数据库、中国 XXX 数据库。

1.3 数据预处理

1）将事故环境数据人工划分为“事故成因”和“事故后果”两类，以便用 ML 工具建立事故与环境要素之间的因果关系。

建立引擎模型的目的是用归纳的方法，试图找出“场景”-“事故类别/事故强度”之间的关联，建立一个以场景要素为输入（原因）、事故类别/事故强度为输出（结果）的分类模型。这个模型主要是用来解决 SOTIF 模型里 Unknown Unsafe（第 III 象限）的问题。

以 NHTSA 的 FARS 数据库为例：

事故记录的要素共有 430 余个，并非所有数据都适合于用来建模。v1.0 按照可感知、可度量的原则，挑选了 81 个参加训练。这 81 个事故描述参数被划分为输入和输出两组。输入组参数是对发生事故时的场景描述，包括人、车、路、设施、天气等要素之间的交互关系；输出组则是事故的类型与事故的严重程度指数。输入参数见下表 1

表 1

Input Module	Element Name (ID (Locator Code)) as per [3]
Scenery Elements	County(C6), City (C7), Month of Crash (C8A), Day of Crash (C8B), Day of Week (C8C), Year of Crash (C8D), Hour of Crash(C9A), Minute of Crash(C9B), Trafficway Identifier (C10), Route Signing (C11), Milepoint (C16), Latitude (C17A), Longitude (C17B), Relation to Junction (C21B), Type of Intersection (C21B), Relation to Trafficway (C23), Work Zone (C24), Light Condition (C25), Atmospheric Conditions (C26), School Bus Related (C27), Rail Grade Crossing Identifier (C28), Body Type (V11), Vehicle Trailing (V14), Gross Vehicle Weight Rating (V18), Vehicle Configuration (V19), Cargo Body Type (V20), Bus Use (V22), Special Use (V23), Emergency Use (V24), Truck Weight Rating (V123), Travel Speed (V25), Driver's ZIP Code (D6), Non-CDL License Type (D7A), Trafficway Description (PC5), Total Lanes in Roadway (PC6), Speed Limit (PC7), Roadway Alignment (PC8), Roadway Grade (PC9), Roadway Surface Type (PC10), Roadway Surface Condition (PC11), Driver's Vision Obscured by (PC14), Marked Crosswalk Present (NM9-PB27), Sidewalk Present (NM9-PB28), X.School Zone (NM9-PB29)
Trigger Event Elements	Critical Event Precrash (PC19), Related Factors- Crash Level (C32), Attempted Avoidance maneuver (PC20), Crash Type-Pedestrian (NM9-PB30), Pre-Event Movement (Prior to Recognition of Critical Event) (PC17), Marked Crosswalk Present (NM9-PB27), Sidewalk Present (NM9-PB28), School Zone (NM9-PB29), Crash Type -Pedestrian (NM9-PB30), Crash Type-Bicycle (NM9-PB30B), Crash Location-Pedestrian (NM9-PB31), Crash Location-Bicycle (NM9-PB31B), Pedestrian Position (NM9-PB32), Bicyclist Position (NM9-PB32B), Pedestrian Initial Direction of Travel (NM9-PB33), Bicyclist Initial Direction of Travel (NM9-PB33B), Motorist Initial Direction of Travel (NM9-PB34), Motorist maneuver (NM9-PB35), Intersection Leg (NM9-PB36), Pedestrian Scenario (NM9-PB37), Crash Group-Pedestrian (NM9-PB38), Crash Group-Cycle (NM9-PB38B)
Driving Status Elements	Related Factors: Driver Level (D24), Sex (P6/NM6), Age (P5/NM5)

输出事故类别根据[20]的分类参数见表 2

CATEGORY	CONFIGURATION	Attribute Codes and Element Attributes
I: SINGLE DRIVER	A: RIGHT ROADSIDE DEPARTURE	1 Drive off Road, 2 Control/Traction Loss, 3 Avoid Collision With Vehicle, Pedestrian, Animal, 4 Specifics Other, 5 Specifics Unknown
	B: LEFT ROADSIDE DEPARTURE	6 Drive off Road, 7 Control/Traction Loss, 8 Avoid Collision With Vehicle, Pedestrian, Animal, 9 Specifics Other, 10 Specifics Unknown
	C: FORWARD IMPACT	11 Parked Vehicle, 12 Stationary Object, 13 Pedestrian/Animal, 14 End Departure, 15 Specifics Other, 16 Specifics Unknown
II: SAME TRAFFICWAY, SAME DIRECTION	D: REAR END	20 Stopped, 21 Stopped, Straight, 22 Stopped, Left, 23 Stopped, Right, 24 Slower, 25 Slower, Going Straight, 26 Slower, Going Left, 27 Slower, Going Right, 28 Decelerating (Slowing), 29 Decelerating (Slowing), Going Straight, 30 Decelerating (Slowing), Going Left, 31 Decelerating (Slowing), Going Right, 32 Specifics Other, 33 Specifics Unknown
	E: FORWARD IMPACT	34 Control/Traction Loss, Avoiding Non-Contact Vehicle- Vehicle's Frontal Area Impacts Another Vehicle, 35 Control/Traction Loss, Avoiding Non-Contact Vehicle- Vehicle Is Impacted by Frontal Area of Another Vehicle, 36 Control/Traction Loss, Avoiding Non-Fixed Object- Vehicle's Frontal Area Impacts Another Vehicle, 37 Control/Traction Loss, Avoiding Non-Fixed Object- Vehicle Is Impacted by Frontal Area of Another Vehicle, 38 Avoiding Non-Contact Vehicle- Vehicle's Frontal Area Impacts Another Vehicle, 39 Avoiding Non-Contact Vehicle- Vehicle Is Impacted by Frontal Area of Another Vehicle, 40 Avoiding Non-Fixed Object- Vehicle's Frontal Area Impacts Another Vehicle, 41 Avoiding Non-Fixed Object- Vehicle Is Impacted by Frontal Area of Another Vehicle, 42 Specifics Other, 43 Specifics Unknown
	F: SIDESWIPE/ANGLE	44 Straight Ahead on Left, 45 Straight Ahead on Left/Right, 46 Changing Lanes to the Right, 47 Changing Lanes to the Left, 48 Specifics Other, 49 Specifics Unknown
III: SAME TRAFFICWAY, OPPOSITE DIRECTION	G: HEAD-ON	50 Lateral Move (Left/Right), 51 Lateral Move (Going Straight), 52 Specifics Other, 53 Specifics Unknown
	H: FORWARD IMPACT	54 Control/Traction Loss, Avoiding Non-Contact Vehicle- Vehicle's Frontal Area Impacts Another Vehicle 55 Control/Traction Loss, Avoiding Non-Contact Vehicle- Vehicle Is Impacted by Frontal Area of Another Vehicle, 56 Control/Traction Loss, Avoiding Non-Fixed Object- Vehicle's Frontal Area Impacts Another Vehicle, 57 Control/Traction Loss, Avoiding Non-Fixed Object- Vehicle Is Impacted by Frontal Area of Another Vehicle, 58 Avoiding Non-Contact Vehicle—Vehicle's Frontal Area Impacts Another Vehicle, 59 Avoiding Non-Contact Vehicle—Vehicle Is Impacted by Frontal Area of Another Vehicle, 60 Avoiding Non-Fixed Object- Vehicle's Frontal Area Impacts Another Vehicle, 61 Avoiding Non-Fixed Object- Vehicle Is Impacted by Frontal Area of Another Vehicle, 62 Specifics Other, 63 Specifics Unknown
	I: SIDESWIPE/ANGLE	64 Lateral Move (Left/Right), 65 Lateral Move (Going Straight), 66 Specifics Other, 67 Specifics Unknown
IV: CHANGING TRAFFICWAY, VEHICLE TURNING	J: TURN ACROSS PATH	68 Initial Opposite Directions (Left/Right), 69 Initial Opposite Directions (Going Straight), 70 Initial Same Directions (Turning Right), 71 Initial Same Directions (Going Straight), 72 Initial Same Directions (Turning Left), 73 Initial Same Directions (Going Straight), 74 Specifics Other 75 Specifics Unknown
	K: TURN INTO PATH	76 Turn Into Same Direction (Turning Left), 77 Turn Into Same Direction (Going Straight), 78 Turn Into Same Direction (Turning Right), 79 Turn Into Same Direction (Going Straight) 80 Turn Into Opposite Directions (Turning Right), 81 Turn Into Opposite Directions (Going Straight), 82 Turn Into Opposite Directions (Turning Left), 83 Turn Into Opposite Directions (Going Straight), 84 Specifics Other, 85 Specifics Unknown
V: INTERSECTING PATHS (VEHICLE DAMAGE)	L: STRAIGHT PATHS	86 Striking From the Right, 87 Struck on the Right, 88 Striking From the Left, 89 Struck on the Left, 90 Specifics Other, 91 Specifics Unknown
VI: MISCELLANEOUS	M: BACKING, ETC.	92 Backing Vehicle, 93 Other Vehicle or Object (2010-2012), 93 Other Vehicle (2013–Later), 98 Other Crash Type, 99 Unknown Crash Type

2）参数归一化处理

将不同年代的不同编码映射成统一编码；将 FARS 和 CRSS 数据合并到一个数据集里。

3）预处理结果

NHTSA_FARS/CRSS 预处理过程文件和结果文件的谷歌云盘链接地址如下：

https://drive.google.com/drive/folders/1d3kElLd_5fj708Si8IaU0DE24f_hv4m?usp=share_link

其他数据库的预处理可以参照上述方法进行。

1.4 生成数据集

数据集是将各个年代、各个数据库的数据统一到一个单一的输入-输出量表格。

模型训练需要区分高速公路与城区工况，所以数据集要分为高速和城区两类。

为生成单一数据集，需要制作数据自动抽取工具，在预处理结果里读值并将其映射成数据集。

以 NHTSA 为例：

数据抽取工具源代码见

https://drive.google.com/drive/folders/1_mnjzfkpIgkKmB7jh8rA2zkGu8ECx9aO?usp=share_link

NHTSA 数据集包括了 1975～2018 年的 1,614,750 个案例，见

https://drive.google.com/drive/folders/1j7J93kDwWFiYHaP1nfBqyHgBhJc49JR2?usp=share_link

1.5 模型学习训练

以 NHTSA_FARS/CRSS 数据库[19]为例。该数据库含有 30 个数据文件，共有大概 430 个场景元素。APE 模型采用了其中的 81 个元素用作模型训练。

1）总体架构

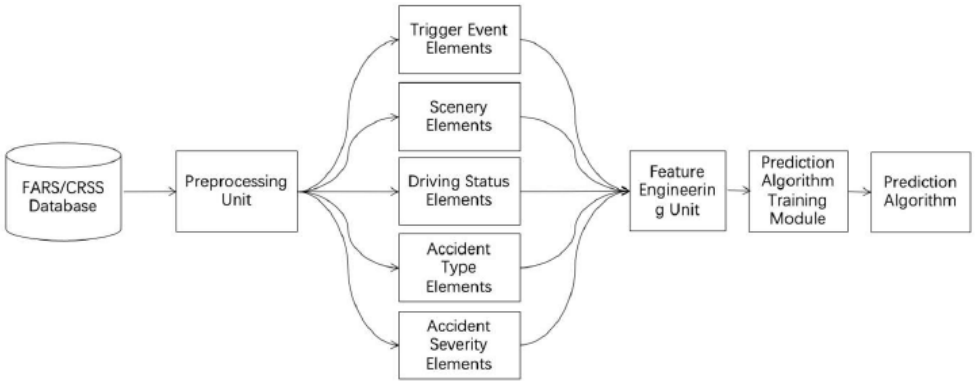

图 48

2）组件构成

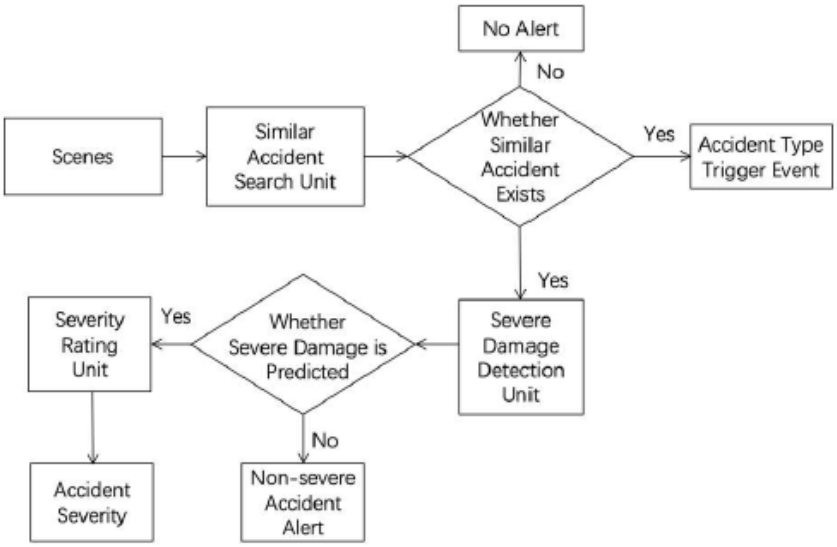

图 49

3）训练流程

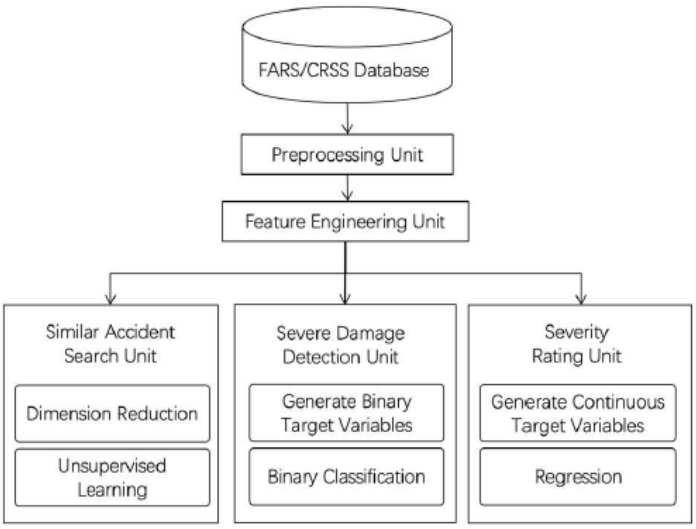

图 50

4）建模结果

建模源代码见

https://drive.google.com/drive/folders/10rPgqQ9l7GK-59YTl_cPdcjA-_E4gOPO

1975～2018 年期间死亡事故的 80%用于训练，剩余 20%用作检验。严重事故类别和严重程度的预测效果分别见图 3、图 4。

	precision	recall	f1-score	support
0	0.82	0.77	0.80	65344
1	0.94	0.96	0.95	257606
accuracy			0.92	322950
macro avg	0.88	0.86	0.87	322950
weighted avg	0.92	0.92	0.92	322950

图 3

	precision	recall	f1-score	support
1	0.78	0.79	0.78	65344
2	0.46	0.53	0.49	77547
3	0.45	0.52	0.49	78375
4	0.36	0.30	0.33	48866
5	0.43	0.29	0.34	52818
accuracy			0.51	322950
macro avg	0.49	0.49	0.49	322950
weighted avg	0.50	0.51	0.50	322950

图 4