Datasheet for '2013-2023 Shooting_datasheet'*

Yanyu Wu

22 April 2024

Motivation

1. *For what purpose was the dataset created?   Was there a specific task in mind?

    -# It appears to address a specific gap by providing a clear and detailed definition of what constitutes a shooting event or firearm discharge, distinguishing these incidents from other forms of violence and thereby offering a focused perspective on gun-related crime.

2. *Who created the dataset (for example, which team, research group) and on behalf of wh

    ion)? * -Toronto Police Services

3. *Who funded the creation of the dataset?   If there is an associated grant, please

-   No

4. *Any other comments?   *

-   No.

Composition

1. *What do the instances that comprise the dataset represent (for example, documents, pho

-   Each row of the dataset is an individual.

2. *How many instances are there in total (of each type, if appropriate)?   *

-   There are 6158 instances.

3. *Does the dataset contain all possible instances or is it a sample (not necessarily ran

---

- A dataset is a sample. The dataset is only about the City of Toronto. The sample is not fully representative of the larger sample set as it covers only part of the geographical area.

4. *What data does each instance consist of?   "Raw" data (for example, unprocessed tex

   -independent

5. *Is there a label or target associated with each instance?   If so, please provide a

- No.

6. *Is any information missing from individual instances?   If so, please provide a des

- No.

7. *Are relationships between individual instances made explicit (for example, users' movi

- No.

8. *Are there recommended data splits (for example, training, development/validation, test

- No.

9. *Are there any errors, sources of noise, or redundancies in the dataset?   If so, pl

- This dataset is updated by hand.

10. *Is the dataset self-contained, or does it link to or otherwise rely on external resour

- This dataset is from the Toronto Police Services https://data.torontopolice.on.ca/datasets/64ddeca12da344

11. *Does the dataset contain data that might be considered confidential (for example, data

- No.

12. *Does the dataset contain data that, if viewed directly, might be offensive, insulting,

- No.

13. *Does the dataset identify any sub-populations (for example, by age, gender)?   If s

- No.

14. *Is it possible to identify individuals (that is, one or more natural persons), either

- No.

15. *Does the dataset contain data that might be considered sensitive in any way (for examp

- No

16. *Any other comments?   *

- No.

Collection process

1. *How was the data associated with each instance acquired?   Was the data directly ob

   -The data associated with each instance is reported by law enforcement based on observ-
   able events and records.

2. *What mechanisms or procedures were used to collect the data (for example, hardware app

- Collection mechanisms include police reports and incident databases, validated through
  law enforcement protocols.

3. *If the dataset is a sample from a larger set, what was the sampling strategy (for exam

- The dataset covers all reported cases in the jurisdiction, but only for Toront..

4. *Who was involved in the data collection process (for example, students, crowdworkers,

   -Data collection was performed by the Toronto Police Services.

5. *Over what timeframe was the data collected?   Does this timeframe match the creatio

   -between 2004 and 2024

6. *Were any ethical review processes conducted (for example, by an institutional review b

- No.

7. *Did you collect the data from the individuals in question directly, or obtain it via t

- Via third parties.

8. *Were the individuals in question notified about the data collection?   If so, pleas

- Yes, it was done by survey.

9. *Did the individuals in question consent to the collection and use of their data?

- No.

10. *If consent was obtained, were the consenting individuals provided with a mechanism to

- No.

11. *Has an analysis of the potential impact of the dataset and its use on data subjects (f

- No.

12. *Any other comments?   *

- No.

Preprocessing/cleaning/labeling

1. *Was any preprocessing/cleaning/labeling of the data done (for example, discretization o

- Yes cleaning of the data was done.

2. *Was the "raw" data saved in addition to the preprocessed/cleaned/labeled data (for exam

- https://data.torontopolice.on.ca/datasets/TorontoPS::shooting-and-firearm-discharges-open-data/about

3. *Is the software that was used to preprocess/clean/label the data available?   If so,

- R was used.

4. *Any other comments?   *

- No.

Uses

1. *Has the dataset been used for any tasks already?   If so, please provide a descriptio

- Yes, Report on Toronto's gun machine trends over the next five years. https://www.cbc.ca/news/canada/to violence-trend-2023-toronto-1.7064901

2. *Is there a repository that links to any or all papers or systems that use the dataset?

- No.

3. *What (other) tasks could the dataset be used for?   *

- This dataset can be used to analyze the relationship between area and casualty data characteristics.

4. *Is there anything about the composition of the dataset or the way it was collected and

- The area of this raw data has deviated from the exact address.

5. *Are there tasks for which the dataset should not be used?   If so, please provide a c

- No.

6. *Any other comments?   *

- No.

Distribution

1.   *Will the dataset be distributed to third parties outside of the entity (for example,

- The dataset is available at opendata Toronto.

2.   *How will the dataset be distributed (for example, tarball on website, API, GitHub)?

- Download from website.

3.   *When will the dataset be distributed?   *

- last refreshed Apr 4, 2024.

4.   *Will the dataset be distributed under a copyright or other intellectual property (IF

- This license is based on the Ontario Open Government License Version 1.0.

5.   *Have any third parties imposed IP-based or other restrictions on the data associated

- No.

6.   *Do any export controls or other regulatory restrictions apply to the dataset or to i

- No.

7.   *Any other comments?   *

- No.

Maintenance

1.   *Who will be supporting/hosting/maintaining the dataset?   *

- No one.

2.   *How can the owner/curator/manager of the dataset be contacted (for example, email ad

- Analytics. Innovation@torontopolice.on.ca

3.   *Is there an erratum?   If so, please provide a link or other access point.*

- No.

4.   *Will the dataset be updated (for example, to correct labeling errors, add new instar

- Quarterly.

5.   `*If the dataset relates to people, are there applicable limits on the retention of th`

- No.

6.   `*Will older versions of the dataset continue to be supported/hosted/maintained?`

- Available on website.

7.   `*If others want to extend/augment/build on/contribute to the dataset, is there a mech`

- freedom. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8383123/

8.   `*Any other comments?        *`

- No.

1 References