**DEEP LEARNING PROJECT PLAN SUMMARY**

* **Research objective**

We will try to classify sounds coming from urban areas. We would like to be able to detect abnormal sounds such as gun shots among more common ones. The idea would be to help the city administration/police detect more quickly dangerous events…

* **Proposed methods**

We have the idea of transforming the sound data into images through the use of ….methods. We will be able to apply convolutional Neural Networks CNN .

* **Dataset**

Music Analysis Dataset : <https://github.com/mdeff/fma>

Million Song Dataset : <http://millionsongdataset.com/>

Open speech : <http://www.openslr.org/12/>

VoxCeleb : <http://www.robots.ox.ac.uk/~vgg/data/voxceleb/>

Urban Sound Classification : <https://datahack.analyticsvidhya.com/contest/practice-problem-urban-sound-classification/>

AudioSet : <https://research.google.com/audioset/?fbclid=IwAR3If9WF29_QwarlvzjwylQVYxxTKNhCAcpA0vanD_hhOe0e8XVfVcyFMYs>

* **Data preprocessing**
* Download the data from the web. Transform the data into images.
* **Estimation method**
* **Hyperparameters tuning**
* **Performence measure**
* **References**
* **Questions to ask**

1. What is the best architecture for sound classification? Is the transformation into images something we should consider doing? Are there any other methods known?
2. Do you see any challenges with our project? What would you like to see in the final report regarding our project?