



Rapport de TP

Module : Apprentissage automatique et réseaux de neurones

Master 1 SII

Mini-projet

Traduction automatique des gestes de la main

- Réalisé par :

BENHADDAD Wissam

BOURAHLA Yasser

23-02-2018

Table des matières

1	Introduction	2
2	Problématique	2
3	Ensemble de données	2
3.1	Description des données	2
3.2	Analyse et prétraitement des données	3
3.2.1	Approche naïve	4
3.2.2	Approche par Clustering	4
4	Solution proposée	5
4.1	Apprentissage sur les données	5
4.2	Application dédiée	5

1 Introduction

L'objectif de ce mini-projet est de nous familiariser avec les concepts et techniques d'apprentissage automatique supervisé (plus particulièrement les réseaux de neurones) afin de mettre en pratique ces aspects théoriques, notamment en utilisant un modèle entraîné sur un ensemble de données pour résoudre un problème en particulier.

2 Problématique

Dans ce projet, nous avons choisis de tenter de réaliser une application qui permettra de traduire différents gestes de la main en un texte ou une action.

Cette traduction automatisée peut servir par exemple à

- Orientation à distance d'un robot.
- Traduction du langage des signes pour faciliter la communication avec les muets.
- **Air-Gesture** Communiquer une action à sa maison, son téléphone ou sa voiture avec la main ...

Il devient évident que réussir à traduire (avec un taux d'exactitude assez raisonnable) des gestes de la main en temps réel (ou différé) s'avère être une tâche irréalisable avec des algorithmes classiques¹, en raison de la complexité de la relation entre les données, cela nous a donc conduit à développer un module d'apprentissage automatique basé sur les réseaux de neurones pour accélérer l'aide à la décision, plus de détails dans MATENSASH TCITER HNA SECTION TA3 NEURAL NEEEETTTTTTTTTTTTTTTTT

3 Ensemble de données

Afin de développer le module d'apprentissage automatique mentionné dans précédemment (voir 2), nous avons choisi le data-set² disponible dans [3], il s'agira donc d'analyser ces données, de les pré-traiter éventuellement afin de les préparer pour la session d'apprentissage CITER FUCKING LEARNING HEEEEERRE

3.1 Description des données

Comme expliqué dans [2], les données d'apprentissage ont été récupérées à l'aide de l'application Vicon-Tracker [1] ainsi que celle de marqueurs (au total de 11) placés sur la face arrière d'un gant, ces derniers servent de source de données envoyées à des capteurs positionnés sur les deux flancs (pour calculer la profondeur), la figure suivante illustre le procédé :

1. Algorithmes naïfs n'ayant pas recours à l'intelligence artificielle pour la résolution de problèmes
2. Ensemble de données d'apprentissage



FIGURE 1 – Environnement ou les données ont été capturées [2]

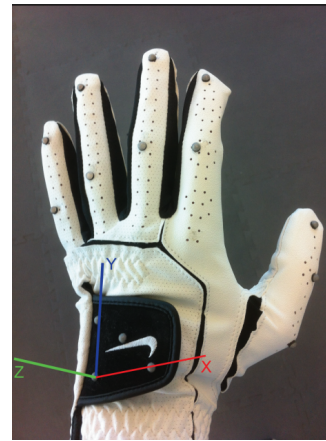


FIGURE 2 – Le gant utilisé pour y attaché les marqueurs(la source des données) [2]

Dans [3], on peut notamment trouvé une brève description des données brutes, elles sont sous forme d'un fichier `.csv` avec le délimiteur de cellules , (virgule), l'entête est structurée de la manière suivante :

- Les deux premières colonnes Class/User représentent respectivement l'identifiant du geste observé (1 à 5) et l'identifiant de l'utilisateur(cobaye) qui a porté le gant durant cette session de collecte de données.
- Les 36 colonnes suivantes contiennent des nombres réels qui représentent les coordonnées cartésiennes en 3-dimensions (x_i, y_i, z_i) des différents marqueurs $Marker_i$.

Il est à noter d'après [2] et [3] que les marqueurs sont non-étiquetés, c'est à dire que pour deux instances I_1 et I_2 du data-set, les coordonnées $(x_i^{I_1}, y_i^{I_1}, z_i^{I_1})$ et $(x_i^{I_2}, y_i^{I_2}, z_i^{I_2})$ ne désignent pas toujours les coordonnées du même marqueurs i .

En raison des conditions de captage des données certaines instances ont des données manquantes représentées par ?

Il est aussi mentionné dans [2] que la présence d'un ou plusieurs marqueurs additionnels est possible(en raison des conditions de captage), en conséquent le nombre de marqueurs visibles varie considérablement selon les instances.

3.2 Analyse et prétraitement des données

Une étape primordiale avant de se lancer dans les essais d'apprentissage est l'analyse et la codification des données, nous avons d'abord effectué une analyse manuelle pour essayer de comprendre comment les données variaient, en nous inspirant des remarques faites dans [2] nous avons remarqué que durant l'enregistrement de la position des marqueurs, certains d'entre eux ne devenaient plus visibles par les capteurs, fortement causé par la nature du geste et non pas à cause d'éventuelles perturbations, pour nous en assurer nous avons dressé le tableau suivant :

		Nombre minimum de marqueurs visibles								
		4	5	6	7	8	9	10	11	12
Geste	1	15604	14190	9191	2639	52	48	48	48	0
	2	14978	14950	14909	14761	14668	14550	13524	10524	31
	3	16322	15380	10746	6423	3556	81	0	0	0
	4	14774	14769	14578	13018	5767	2349	13	0	0
	5	15727	15686	15648	15406	14900	13535	10382	4180	0

TABLE 1 – Nombre d’instances par classe de geste dont le nombre minimum de marqueurs est visible (donnée non manquante dans le data-set)

Nous pouvons observer que par exemple, pour le geste 2 **tous les doigts levés** chaque instance étiquetée par cette classe possède quasiment tout les marqueurs visible à l’exception du dernier(comme mentionné dans 3.1 ce marqueurs est probablement en plus), il en est presque de même pour le geste 5 (**geste de saisie un sorte de poing ouvert légèrement**), nous pouvons voir que qu’il y a une petite différence avec les données de la classe 2, cela est du au fait que les marqueurs placés aux ongles ne sont pas toujours visibles.

Pour les classes 3 et 4, elles diffèrent en un seul marqueurs ou deux car les deux sont très similaires(3 : **un doigt levé**, 4 : **deux doigts levés**), enfin pour les geste de la classe 1 (**Poing fermé**) le manque apparent de marqueurs visibles est expliqué par le fait que les marqueurs au jointure et extrémité des doigts soient masqués par la nature du geste en lui même.

Pour l’ensemble des classes, il es très rare qu’on trouve au minimum 12 marqueurs visibles a cause du bruit présent dans les données(voir 3.1).

Bien que les valeurs manquantes ont une grande signification pour nous humains, il faut tout de même traduire cette information pour le modèle, deux approches ont été pensée.

3.2.1 Approche naïve

Affecter une(des) valeur(es) aux données manquantes est une tâche souvent ardue, une multitude de théories existent sur le sujet, mais parfois la solution la plus simple semble être la plus efficace, nous avons simplement remplacé le valeurs manquantes par des **zéros**, ce choix est appuyé par un raisonnement fait à partir de l’analyse faite préalablement, en effet chaque classe de geste suit un certain motif(pattern) de valeurs manquantes, autrement dit si nous devons remplacer ses valeurs manquante par une approximation(une moyenne sur les valeurs de la colonne par exemple) cela pourrait rapprocher les instances entre elles et donc pourrait fausser l’apprentissage, car les valeurs des attributs seraient trop proches l’une d’entre elles, les remplacer par une valeur nulle pourra en théorie aider le modèle a mieux distinguer les classes, pour ainsi mieux approximer le pattern que suivent ces données.

3.2.2 Approche par Clustering

4 Solution proposée

4.1 Apprentissage sur les données

4.2 Application dédiée

Table des figures

1	Environnement ou les données on été capturées [2]	3
2	Le gant utilisé pour y attaché les marqueurs(la source des données) [2]	3

Liste des tableaux

1	Nombre d’instances par classe de geste dont le nombre minimum de marqueurs est visible (donnée non manquante dans le data-set)	4
---	---	---

Références

- [1] Vicon datastream sdk developer’s manual, Jan 2013.
- [2] A. Gardner. Datasets for motion-capture-based hand gesture recognition. <http://www2.latech.edu/~jkanno/datadescription-1.pdf>, May 2017.
- [3] A. Gardner, R. R. Selmic, and J. Kanno. Mocap hand postures (2016). <https://archive.ics.uci.edu/ml/datasets/MoCap+Hand+Postures>, Nov 2016.