Identifying the Sport Activity of GPS Tracks

Cèsar Ferri

DSIC, Universitat Politècnica de València
Camí de Vera s/n, 46022 València, Spain.
cferri@dsic.upv.es

Abstract

The wide propagation of devices, such as mobile phones, that include a global positioning system (GPS) sensor has popularised the storing of geographic information for different kinds of activities, many of them recreational, such as sport. Extracting and learning knowledge from GPS data can provide useful geographic information that can be used for the design of novel applications. In this paper we address the problem of identifying the sport from a GPS track that is recorded during a sport session. For that purpose, we store 8500 GPS tracks from ten different kinds of sports. We extract twelve features that are able to represent the activity that was recorded in a GPS track. From these features several models are induced by diverse machine learning classification techniques. We study the problem from two different perspectives: flat classification, i.e, models classify the track in one of the ten possible sport types; and hierarchical classification, i.e. given the high number of classes and the structure of the problem, we induce a hierarchy in the classes and we address the problem as a hierarchical classification problem. For this second framework, we analyse three different approaches. According to our results, multiclassifier systems based on decision trees obtain the better performance in both scenarios.

Keywords: Classification, Hierarchical Classification,Geographic Information, GPS

1 Introduction

The last generation of mobile phones, as well as other electronic devices, include different kinds of sensors able to capture the users real-time situational information, such as location, speed, etc. One of the most powerful sensors is the GPS receiver that allows users to locate spatially the device and track its position regularly. This data provides a source of innovation in producing innovative applications to the device user. Recording GPS data for its analysis can be very useful in different fields. Specifically, the use of GPS tracks in machine learning has been employed for different purposes. Some works use GPS data for several goals in health sciences [1] and agronomic sciences [13]. Recognising the transportation modes of people have important applications in pervasive computing [18].

Since the increase in popularity of GPS devices, they are widely used for tracking sport activities. For this purpose, there exist popular apps such as *Runtastic* or *Endomondo* that

allows users to register the tracks of the activities and then the user can share these tracks in specialised social networks. A popular example of this kind of social networks is Wikiloc[1]. In this social network more than 1,000,000 users share about 2,500,000 GPS tracks of about 72 different kinds of sports such as Hiking, Cycling, Sailing, Snowshoeing.... These GPS tracks are located all around the world. Other examples of social networks based on GPS tracks are: *MapMyTracks*[2] and *everytrail*[3]. In these social networks every time a user uploads a new track, he must manually introduce the kind of sport activity for the GPS track.

In this paper we investigate how to identify automatically the kind of sport during the activity that generated the GPS track with machine learning techniques. For that purpose, we downloaded 10,000 GPS tracks of ten popular sports: Running, Cycling, Hiking... Using a GPS library, we extracted a set of twelve features such as duration, length, average and maximum speed.

We address the problem of sport identification from two different perspectives. First we consider this problem as a traditional "flat" classification problem. In this way we measure the performance of twelve different classification methods in predicting the sport that generated the GPS from the ten possible alternatives. Given the high number of classes existing in the problem we also study the performance of the methods when we build a tree of specialised local classifiers.

The second perspective is based on the inherent hierarchical structure of the classes of the problem. For instance, tracks generated when the user is using a bike (Cycling or Mountain Biking) should have more similar features (length, average speed..) with respect to tracks generated by aquatic sports (Kayaking and Sailing). In this case, many real-world classification problems are naturally cast as hierarchical classification problems [14]. In this kind of problem, the classes to be predicted are organised into a class hierarchy, usually a tree. According to [14] there are three main different broad types of approach: flat, i.e., ignoring the class hierarchy; local, the hierarchy is taken into account by using a local information perspective, i.e. building local and specialised classifiers; and global, a single classification model is built taking into account the class hierarchy. Additionally, a hierarchical classification method can be implemented in a way that the learned model will always classify a leaf node (mandatory leaf-node prediction) or the model can also consider stopping the classification at any node in any level of the hierarchy (non-mandatory leaf-node prediction).

Considering the task of identifying a sport from a GPS track as a hierarchical classification problem, we first induce a class hierarchy tree based on the features of the sport and common misclassification errors between classes. Using this class hierarchy tree of sports, we compare different approaches with a distance-based performance metric for hierarchical classification problems. Specifically, we analyse flat classification; specialised local classifiers in a mandatory leaf-node prediction context; a variation of the local approach where we add rejection rules [10] in order to predict classes at any level of the hierarchy; and finally, a method for labelling of instances based on the minimisation of the expected loss.

This paper is structured as follows. Section 2 includes details about the process of collecting GPS tracks and extracting features. The techniques used for addressing the problem as flat classification are explained in Section 3. Section 4 details the methods and results obtained when we consider the problem as hierarchical classification. A summary of related works is included in Section 5. We finish the paper with the conclusions in Section 6.

---

[1] http://www.wikiloc.com/wikiloc/home.do
[2] https://www.mapmytracks.com/
[3] http://www.everytrail.com/

Feature	Description
length_2d	Length (2-dimensional) of route in kilometres
length_3d	Length (3-dimensional) of route in kilo-meters
moving_time	Time (seconds) of track with significant movement
stopped_time	Time (seconds) of track without no significant movement
moving_distance	Distance (meters) travelled during stopped times
stopped_distance	Distance (meters) travelled during stopped times
max_speed	Maximum speed (m/s) of the track
ave_speed	Average speed (m/s) of the track
uphill_norm	Uphill elevation climbs divided by length_2d, in meters/km
downhill_norm	Downhill elevation climbs divided by length_2d, in meters/km
max_elev	Maximum elevation in meters of the track
min_elev	Minimum elevation in meters of the track

Table 1: Features extracted from a GPS track.

2 Data Collection and Feature Extraction

In this section we give details about the recollection and manipulation of the GPS data.

2.1 GPS Tracks

GPS devices usually store tracks as GPX files. GPX (GPS Exchange Format) is an XML schema designed as a common GPS data format for software applications. It can store waypoints, tracks, and routes. In our case, we work with GPX files containing a track. This track is made of at least one segment containing waypoints, that is, an ordered list of trackpoints describing a path. Each trackpoint contains position (latitude and longitude), elevation and a timestamp. Latitude and longitude are expressed in decimal degrees. Elevation is recorded in meters. Dates and times are Coordinated Universal Time (UTC).

2.2 Feature Extraction

In order to identify the sport that generated the GPS track, we need to create a set of features able to describe accurately the activity. These features will then be useful to determine the kind of sport of a GPS track. For instance, if we detect a high speed segment in the GPS track, it logically must be created practising a sport that involves the use of bicycles or motorbikes.

There exist software libraries that can analyse a GPS track and produce a set of features that describe the track contained in the GPX file. In this work we use *gpxpy*[4] a simple python library for parsing and manipulating GPX files. Table 1 contains the features that we extract for each GPS track. Most of them are directly or indirectly provided by the *gpxpy* library. We have incorporated three features that we consider important for our purpose: Average speed, Maximum elevation and Minimum elevation.

2.3 Dataset

In this work, we concentrate on the most popular sports according to the number of tracks uploaded to social networks. Concretely, we selected ten different kinds of sports. In Table 2 we include the sports selected, its acronym and a brief description.

---

[4]Library created by Tomo Krajina, https://github.com/tkrajina/gpxpy

Acr	Sport	Description
rn	Running	The sport of someone who runs
hi	Hiking	Sport consisting of vigorous walks, usually on trails off-road
mb	Mountain biking	Sport of riding mountain bikes usually off-road
ci	Cycling	Bicycle racing sport usually held on paved roads
tr	Trail running	Sport consisting of running over trails
mo	Mountaineering	Sport of mountain climbing
mc	Motorcycling	Sport of travelling on a motorcycle
tb	Trail bike	Recreational off-road and on-road riding of motorcycles
ky	Kayaking	Sport consisting in the use of a kayak for moving across water
sl	Sailing	Sport of riding in a sailboat

Table 2: List of sports, acronyms, and a brief description.

Note that in some cases it is difficult to differentiate clearly between some sports. For instance, a hiking GPS track uploaded by an expert hiker can be very similar to the track uploaded by a beginner trail runner. We find a similar situation between cycling and mountain biking sessions. Here the main difference is the kind of bike used in the activity as well as the kind of road, but these features are not directly represented in a GPS track. With this in mind, we studied the possibility of adding information about the the kind of track surface (on-road, off-road, sea, river). There are some web services that offer limited information such as knowing if some GPS coordinates are located in the sea, however their use is very restricted. Therefore we decided against incorporating this information into the dataset.

After the list of sports was selected, we downloaded 1,000 GPS tracks for each of the ten selected sports from a sport social network. Every GPS track was labelled manually by the users according to the sport they practised. In order to exclude fake or anomalous GPS tracks, we discarded the tracks that do not satisfy a set of requirements of length of route and time of track. After this filtering process, we randomly selected 850 tracks for each type of sport to produce a balanced dataset of 8,500 GPS tracks. Each track is characterised by the twelve features included in Table 1. Figure 1 contains histograms for these twelve features. These plots can help us to know the variability of the features in the dataset. An average activity implies covering a distance of 44.22 km in 2 hours and 53 minutes with a maximum speed of 5 m/s (18 Km/h).

In Table 3 we analyse the averages of all the features for the ten kinds of sports. As expected, we find important differences among the classes. The sport with the shortest activity in distance and time is running, while Motorcycling is the longest in distance and Mountaineering the longest in time. If we consider speed, Mountaineering is the slowest activity and Motorcycling is the fastest sport. Mountaineering is also the sport that reaches maximum elevations, and sailing is logically the closest activity to the sea level, although it seems that some of the sailing sessions have been performed in lakes. We believe that the high values of *downhill_norm* and *uphill_norm* features in sailing activities are due to poor calibration of the GPS receivers.

3   Identifying the Sport by Classification

After the collection of tracks and the feature extraction phase, we created a dataset formed by 8500 tracks, 12 attributes and ten different classes. From this dataset we can apply classification techniques in order to know if they are able to predict accurately the sport that generated a GPS track. In the experiments, we used some of the classification methods of Weka suite in a R script
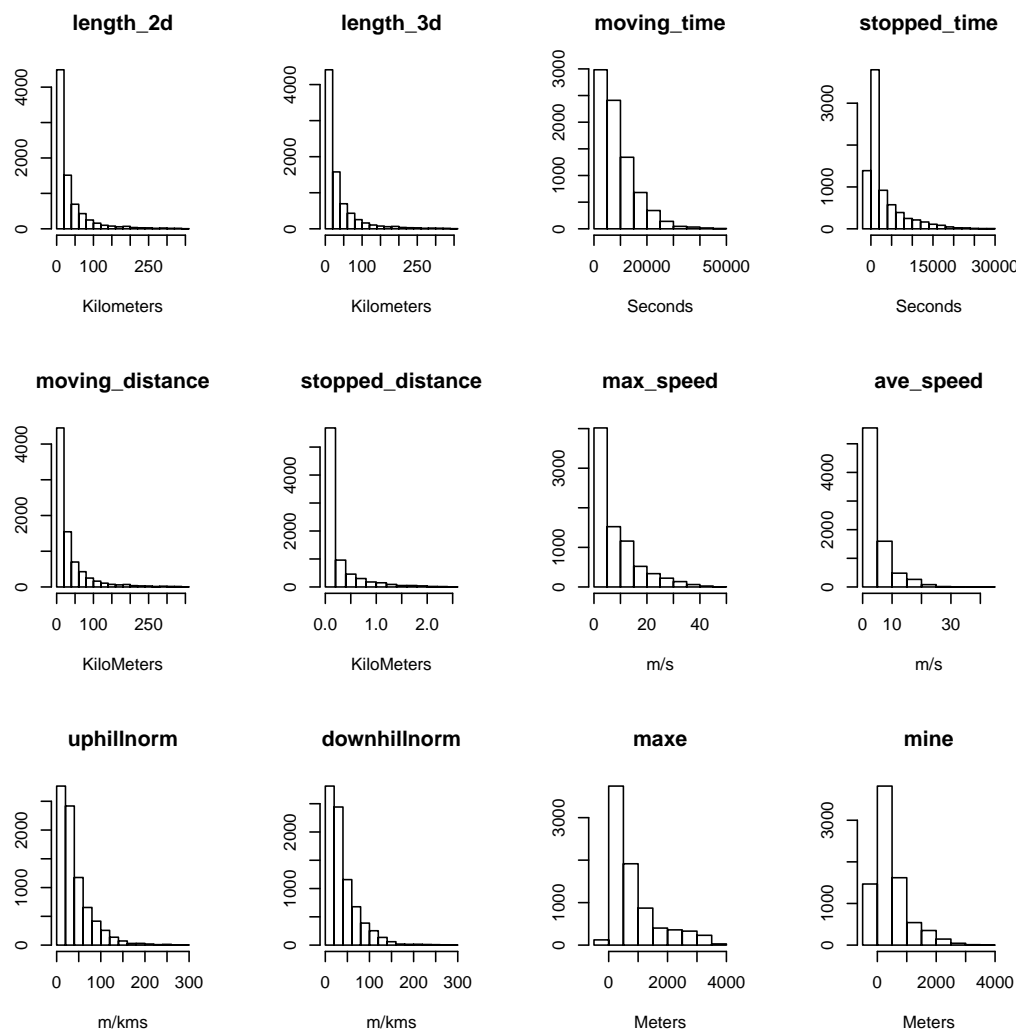
Figure 1: Histograms of the twelve features of the GPS track dataset.

by means of the library RWeka and caret . Specifically, we use the following twelve classification methods: a decision tree "J48", a propositional rule learner "JRip", logistic regression "Logist", naive Bayes "NB", K-nearest neighbours with ten neighbours "IBK", Random Forest "RF", a combination of ten J48 models by the Bagging technique "Bagging", a decision list "PART", a combination of ten J48 models by the Boosting technique "Boosting", a support vector machine "SVM" , a Boosted Logistic Regression "LB" and a Mixture Discriminant Analysis model "MDA". In all the methods we used the default parameters except that we assigned the parameter k to ten in "IBK", and the selection of "J48" as base classifier in Boosting and Bagging. The results in accuracy of all these methods are shown in Table 5. These results are the average of executing ten times a ten fold cross validation evaluation.

Feat.	len_2d	len_3d	movtime	stptime	mvdist	stpdist	maxsp	avesp	uphill	dwhill	maxe	mine
Units	km	km	sec.	sec.	km	km	m/s	m/s	m/km	m/s	met.	met.
rn	8.68	8.85	3237.21	189.06	8.82	0.02	4.72	2.94	36.03	35.45	384	289
tr	16.47	16.94	7935.81	742.24	16.80	0.08	4.09	2.34	57.01	55.05	827	409
hi	12.06	12.48	9410.45	3059.37	12.17	0.28	2.79	1.47	65.23	60.50	948	574
mo	13.96	15.01	14395.95	8399.82	14.13	0.83	2.51	1.23	110.04	97.94	2464	1426
ky	14.73	15.02	7318.30	3017.64	14.82	0.17	4.40	2.62	26.34	26.65	198	141
sl	35.19	35.43	9075.64	2579.03	35.24	0.18	7.85	4.91	18.22	17.93	101	49
mb	36.10	36.46	9382.89	2094.88	36.22	0.14	9.20	4.58	27.33	27.64	688	346
ci	54.36	54.66	9428.22	1589.04	54.53	0.10	10.34	5.69	22.95	22.63	674	342
mc	111.11	111.52	8335.79	3594.34	111.28	0.23	23.18	12.97	19.20	19.45	953	370
tb	66.64	67.12	8723.57	3611.26	66.87	0.23	16.46	8.25	28.20	27.73	911	416

Table 3: Averages of the twelve features disaggregated by the kind of sport.

Class	Predicted Class										Perf. Metrics by class		
	ci	hi	ky	mb	mc	mo	rn	sl	tb	tr	Prec	Recall	F-Meas
ci	**512**	4	10	182	21	2	20	32	58	9	0.631	0.602	0.616
hi	4	**608**	54	8	6	95	21	6	8	40	0.663	0.715	0.688
ky	19	87	**555**	18	10	3	29	98	21	10	0.707	0.653	0.679
mb	103	6	3	**653**	9	3	8	18	27	20	0.675	0.768	0.718
mc	30	2	7	14	**607**	2	9	18	155	6	0.686	0.714	0.700
mo	2	85	6	6	3	**716**	6	1	4	21	0.834	0.842	0.838
rn	20	23	39	11	10	5	**584**	29	6	123	0.675	0.687	0.681
sl	32	12	89	19	23	3	40	**603**	19	10	0.724	0.709	0.717
tb	75	20	9	39	192	9	15	23	**446**	22	0.585	0.525	0.553
tr	14	70	13	18	4	21	133	5	18	**554**	0.680	0.652	0.665
Mean											0.686	0.687	0.686

Table 4: Confusion matrix for Bagging ten J48 trees (ten fold cross validation evaluation). On the right part of the table we include performance metrics by class.

3.1 Flat Classification Approach

In Table 4 we include the confusion matrix for a ten fold cross validation evaluation of Bagging ten J48 trees. This confusion matrix shows how the classifier distributes the errors. We also include detailed information of some performance metrics by class [7]. The sport that is better predicted is Mountaineering, since we reach a partial accuracy of 84% of the cases. The most difficult sport is Trail bike mainly because the classifier assigns many of the Trail bike tracks to Motorcycling tracks. In fact, the confusion matrix expresses a quite predictable behaviour in the sense that we can see misclassifications between similar sports such as those mentioned previously. Other examples of this behaviour are: Cycling-Mountain biking, Sailing-Kayaking, Hiking-Mountaineering and Running-Trail running.

3.2 Local Classifier Approach

When we address a problem with a relatively high number of classes, there exists the possibility of decomposing the global classification problem into smaller problems by grouping similar classes, and thus forming a hierarchy of classes. This approach is named local classifier approach. In this work, we use the Local Classifier Per Parent Node schema (terminology of [14]). Starting from a given hierarchy of classes (where the original classes are placed in the leaves of the tree), we learn a specialised classifier for each branch node of the hierarchy tree, i.e. a local specialised classifier. A global classifier loses the intuition that classes that are close to each other in the hierarchy have more similarities with each other, in general, than topics that are far apart in the hierarchy. In our case, it is less difficult to classify that a track has been generated by a running activity or a walking activity than to learn a classifier that is able to predict the
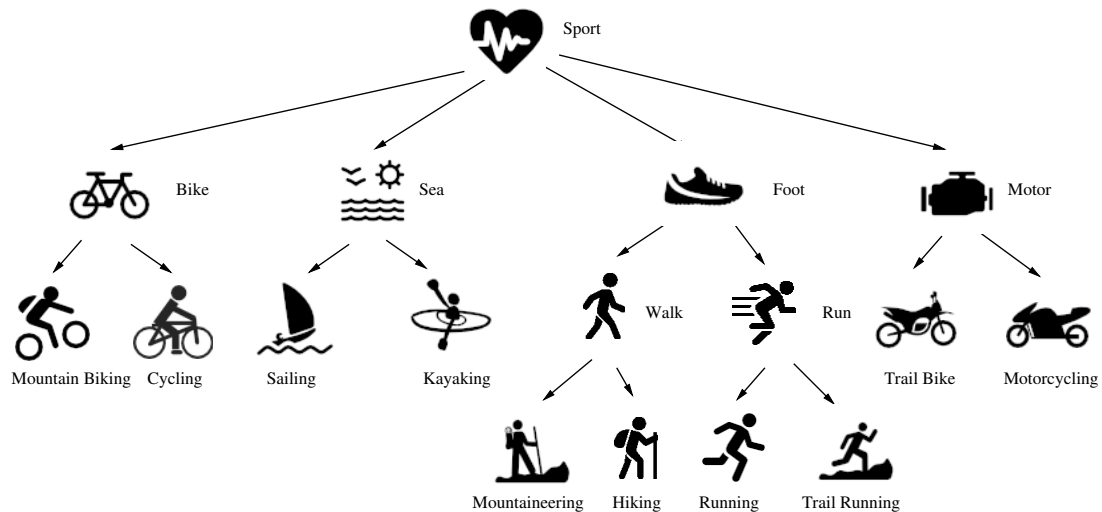
Figure 2: Class hierarchy tree. Original classes correspond to the leaf nodes.

ten kinds of sport activities correctly. A drawback of the top-down class-prediction approach is that an error at a top class level is going to be propagated down the hierarchy. There are some techniques that try to reduce this problem, some of them by improving class probability estimations like shrinkage [9] and isotonic smoothing [11].

Some studies present methods to build class hierarchies from the similarity of classes within data [8]. In our case, we propose the hierarchy of Figure 2. Given the properties of the problem in question, it is not difficult to induce a class hierarchy tree by using the similarities among the ten sports that the data reveals. In our proposal, we use locomotion form, surface type and speed in order to cluster sports. Although there are other alternatives, for instance join Hiking with Running, and Trail running with Mountaineering, if we analyse Table 4 we can see that the joined classes are in most cases the ones with more common misclassifications.

Therefore, following the tree of classes[5] of Figure 2, we inferred seven different local classifiers corresponding to the seven branch nodes of the tree. Each classifier is specialised in classifying between the class siblings of the branch. In order to learn each classifier, we joined the corresponding classes following the tree. In Table 5 we can see the results in accuracy of this proposal in comparison with the same learning techniques employed in the flat classification approach. These results correspond to the averages of executing ten times a ten fold cross validation evaluation. In general, if we compare the flat classification approach with the local classification approach, we see that there is not a general pattern. For some techniques (especially JRip) the local approach is able to improve the results of the traditional flat methodology, however we also find techniques with the inverse behaviour (Logist, NB...). If we only consider the best methods (multiclassifiers) the differences are small.

4 Hierarchical Classification Approach

Given the hierarchy inherent in the classes, we analyse this problem as a hierarchical classification problem using for that purpose the tree of Figure 2. For this framework, we consider that

---

[5]Icons made by Freepik and icons8 from www.flaticon.com and licensed by Creative Commons.

	J48	JRip	Logist	NB	IBK	RF	Bagg.	PART	Boost.	SMV	LB	MDA
Flat	0.616	0.564	**0.569**	**0.455**	0.558	**0.677**	**0.669**	0.605	0.669	**0.541**	**0.611**	**0.508**
Local	**0.628**	**0.609**	0.548	0.415	**0.564**	0.676	**0.669**	**0.610**	**0.671**	0.530	0.592	0.504

Table 5: Accuracy of the classification methods (average of ten times a ten fold cross validation). Flat and local approaches. We highlight in bold the best result for each learning method.

hierarchical classification can label a new instance with a label belonging to one of the branch nodes of the class hierarchy tree. This could occur when the hierarchical classifier determines that it is not reliable enough to descend in the hierarchy and return a leaf node class. Following the terminology of [14] this corresponds to a non-mandatory leaf node prediction approach. On the contrary, the classifier approaches of Section 3 must always return a class placed in a leaf node of the tree (a mandatory leaf node prediction in the terminology of [14]). For our problem, identifying the sport activity that is generated by a GPS track, this hierarchical approach can be useful for the situations when the model is not sure about the confidence of the predicted class. In these cases the model could stop in a non-leaf class, and then it could suggest the leaf classes descending from that node branch to the user.

For the hierarchical classification approach we cannot use accuracy as a performance measure since all the train instances belong to leaf classes, while some predictions will be branch classes. Several performance metrics have been defined specifically for the hierarchical classification task. A review of these measures can be found in [3]. In this work we use a metric based on the distance in the tree of classes between the predicted and the actual class. Concretely, we use a metric (hd) defined as the number of edges that traverses the shortest path between the predicted and the actual class (also known as tree-error [4]). We divide this amount by the number of edges of the longest path in the trees in order to obtain a quantity between 0 and 1. For instance, if the prediction is Mountain Biking and the real class is Trail Bike, we have that the path between these to classes traverses 4 edges. Given that the longest path in the tree has 5 edges $hd(mb,tb)=4/5$. This measure is based on the rationale that classes that are close to each other in the tree of hierarchies tend to be more similar to each other than other classes. A similar evaluation metric was used in [15].

4.1 Non-Mandatory Leaf Node Prediction based on Probability Estimations

Here, we introduce a technique that is able to return classes that are not in the leaves by using flat classifiers and the tree of hierarchies. The idea is using the scores computed by flat classifiers as probability class estimations. With these estimations and the distances from the tree of hierarchies we can compute the expected hd value for all classes and then we select the class with the lowest expected value. This is similar to labelling instances that minimise the expected misclassification cost in cost-sensitive classification. Formally, we define C as the set of original classes (in our problem the ten different sports) that correspond to the classes in the leaves of the tree, we define C_{ext} as C plus the classes in the branches of the tree of classes (in our problem $Sport, Bike, Foot...$), $p(i,c)$ returns the probablity estimation that an example i belongs to a class $c \in C$, and $hd(c_i, c_k)$ computes the hd measure for $c_i, c_j \in C_{ext}$. We define the function $hprb$ for classifying instances as: $hprb(i) = argmin_{c \in C_{ext}}(\sum_{k, \forall k \in C} p(i,c) * hd(k,c))$.

4.2   Non-Mandatory Leaf Node Prediction based on Thresholds

When we are working with local classifiers, a direct way to deal with the non-mandatory leaf-node prediction problem is to use a "blocking criterion" based on thresholds at each class node. In this way, the classification stops for an instance if the confidence score of the classifier at a given class node for that instance is lower than a defined threshold. A method for automatically computing these thresholds is introduced in [2].

Here, we use a simple abstaining method inspired by the delegating classifiers proposed in [6]. Given a confidence parameter t, and a function $p_n(i, c)$ that returns the probably estimation that an example i belongs to a class $c \in C_n$, where C_n is the set of classes of a node n of the tree, we define a stopping criterion that halts the descending process in a node n if $max(p_n(i, c) < ((1/|C_n|) + t))$. In this way, we return the node class if any of the estimated probabilities is bigger than the confidence threshold (computed as $(1/|C_n|) + t$). Note that the confidence threshold depends on the number of classes C_n. When we set $t = 0.3$, if $|C_n| = 2$, then the confidence threshold will be 0.8, but if $|C_n| = 4$ the confidence threshold will be 0.55. In our experiments we set a confidence parameter $t=0.3$.

In Table 6 we can see the Mean squared error (MSE) of the twelve methods and the results in hd of the flat classifiers detailed in Section 3.1 and the models induced by the local classifier approach 3.2. We also show the hd of the $hprb$ approach and the variation of the local classifiers approach with a blocking criterion based on thresholds. First, we see that in general Non-Mandatory Leaf Node Prediction approaches ($hprb$ and $Local + block$) present better performance than Mandatory Leaf Node Prediction approaches ($Flat$ and $Local$) according to results in hd. The only exception is SVM where the best method is the flat approach. This result is probably caused by the poor quality of the probability estimations in this method (reflected in the high value of MSE). If we study the insertion of the block strategy in the local classifier approach with respect to the original method, we see that this modification allows us to improve in all the cases the performance metric except from SVM, and therefore shows the adequacy of stopping the descending of the tree of classes when there is uncertainty. Finally, comparing Non-Mandatory Leaf Node Prediction approaches ($hprb$ and $Local + block$), we can observe that $hprb$ obtains better results in seven methods, while $Local + block$ gets the best results in four techniques. According to the MSE of the methods, it seems that good probability estimations are more useful for $Local + block$ since the methods with low values of MSE present the best results with the $Local + block$ approach. Our intuition is that bad estimations of probabilities can cause bad decisions in the top levels of the classifiers tree and this damages the final performance of the $Local + block$ approach.

5   Related Work

In pervasive computing, it is an important research problem to recognise the transportation modes of people. Several works have addressed this problem. A proposal to automatically learn transportation mode from raw GPS data is presented in [18]. Four different inference models (DTs, Bayesian Nets, SVMs and CRFs) are analysed in the experiments. The decision tree model obtains the better results in terms of accuracy. The data is collected using the GPS data of 45 users over six months period. The approach of [17] consists in detecting transportation mode recognition on mobile phones only using the embedded accelerometer. A similar work is [5]. In this research, the authors collected 150 hours of GPS and accelerometer data from two users that practise five activities: bicycling, walking, riding in a vehicle, sitting, and standing. The authors extract 49 features of this data and compare the performance of several machine

	MSE	Flat	Hprb	Local	Local+block
J48	0.0580	0.2571	0.2567	0.2473	**0.2449**
JRip	0.1305	0.3179	0.2789	0.2684	**0.2670**
Logist	0.1837	0.2963	**0.2844**	0.3184	0.3053
NB	0.0737	0.4009	**0.3913**	0.4200	0.4091
IBK	0.1574	0.3036	**0.2730**	0.2995	0.2857
RF	0.1300	0.2116	**0.1988**	0.2113	0.2033
Bagg.	0.1228	0.2166	**0.2037**	0.2165	0.2095
PART	0.0545	0.2661	0.2655	0.2594	**0.2533**
Boost.	0.0278	0.2147	0.2117	0.2123	**0.2098**
SMV	0.2723	**0.3221**	0.4296	0.3275	0.4800
LB	0.1468	0.2619	**0.2438**	0.2749	0.2599
MDA	0.1720	0.3477	**0.3320**	0.3540	0.3368

Table 6: MSE of the twelve methods and *hd* obtained by using two Mandatory Leaf Node Prediction approaches (*Flat* and *Local*) and two Non-Mandatory Leaf Node Prediction approaches (*hprb* and *Local + block*). The results show the average of ten times ten fold cross validation evaluation. The best approach in *hd* for each learning method is in bold.

learning methods. Only 6 of these features are obtained from the GPS data: average speed, net distance travelled, and four more related to the GPS signal quality. In this work random forests is the technique with the better performance.

An interesting paper is [16]. The goal of this work is to predict different activity modes from the combination of GPS and accelerometer data. In this case ten volunteers wore the sensors during different sessions where they practised these activities: walking, jogging, bicycling, inline skating, or driving an automobile. The authors selected a set of features from accelerometer counts, and steps and GPS speed by a discriminant function analysis. The best performance is obtained by using three variables from the accelerometer and three from the GPS (median counts, steps and speed). The authors claim that this pilot study provides evidence that the use of GPS together with the accelerometer improves physical activity mode classification to a small degree, although they find that larger studies among free-living individuals and with an expanded range of activities are required to assess this conclusion.

There are also some articles that study the use of GPS data for several goals in social sciences and medicine. A summary of different works can be found in [1]. The purpose of this paper is to review the utility of the GPS data in the study of health-related physical activity. The paper remarks that GPS, especially when used in combination with GIS and accelerometers, is a powerful tool for studying the relationship of environmental attributes to human behaviour in terms of physical and transport-related activities. Finally, in [12] the authors try to identify the proportion of children's physical activity occurring in public parks with playgrounds.

6   Conclusions

Recognising the transportation modes of people is an open research problem with applications in fields such as pervasive computing and mobile applications. In this paper, we aim to identify the sport activity related to a GPS track. For that goal, we compiled a set of features from a GPX file and we studied the predictive performance of several machine learning methods. We gathered 8500 GPS tracks from social networks of ten kind of sport activities. In this aspect, this work addresses a more complex and realistic scenario when we consider related work. We have analysed the problem from two perspectives: flat classification and hierarchical classification. In this second case we compared different hierarchical classification approaches.

The employment of these models can provide knowledge and information for the construction of novel applications and the improvement of user experience in different ways. For instance, if a mobile device is able to detect when a user has started a running session, it could automatically set a "running mode" with a specific configuration adapted to that context (noise level, phone interface...).

As future work we propose the integration of surface information features (on-road, off-road, river..) that could probably improve the prediction performance. In this direction, the integration of data collected from the accelerometer of the device could also be useful, as related works have shown.

Acknowledgments

References

[1] Nancy F Butte, Ulf Ekelund, and Klaas R Westerterp. Assessing physical activity using wearable monitors: measures of physical activity. *Medicine and science in sports and exercise*, 44(1):5–12, 2012.

[2] Michelangelo Ceci and Donato Malerba. Classifying web documents in a hierarchy of categories: A comprehensive study. *J. Intell. Inf. Syst.*, 28(1):37–78, 2007.

[3] E Costa, A Lorena, ACPLF Carvalho, and A Freitas. A review of performance evaluation measures for hierarchical classifiers. In *Evaluation Methods for Machine Learning II: papers from the AAAI-2007 Workshop*, pages 1–6, 2007.

[4] Ofer Dekel, Joseph Keshet, and Yoram Singer. Large margin hierarchical classification. In *Proceedings of the twenty-first international conference on Machine learning*, page 27. ACM, 2004.

[5] K Ellis, S Godbole, S Marshall, G Lanckriet, J Staudenmayer, and J Kerr. Identifying active travel behaviors in challenging environments using GPS, accelerometers, and machine learning algorithms. *Frontiers in Public Health*, pages 2–36, 2014.

[6] César Ferri, Peter Flach, and José Hernández-Orallo. Delegating classifiers. In *in Proc. 21st International Conference on Machine Learning*, pages 27–36. Omnipress, 2004.

[7] César Ferri, José Hernández-Orallo, and R Modroiu. An experimental comparison of performance measures for classification. *Pattern Recognition Letters*, 30:27–38, 2009.

[8] Tao Li, Shenghuo Zhu, and Mitsunori Ogihara. Hierarchical document classification using automatically generated hierarchy. *Journal of Intelligent Information Systems*, 29(2):211–230, 2007.

[9] Andrew McCallum, Ronald Rosenfeld, Tom M Mitchell, and Andrew Y Ng. Improving text classification by shrinkage in a hierarchy of classes. In *ICML*, volume 98, pages 359–367, 1998.

[10] Tadeusz Pietraszek. Optimizing abstaining classifiers using roc analysis. In *Proc. of the 22nd International Conference on Machine learning*, pages 665–672. ACM, 2005.

[11] Kunal Punera and Joydeep Ghosh. Enhanced hierarchical classification via isotonic smoothing. In *Proceedings of the 17th International Conference on World Wide Web*, pages 151–160. ACM, 2008.

[12] Robin Quigg, Andrew Gray, Anthony I Reeder, Alec Holt, and Debra L Waters. Using accelerometers and GPS units to identify the proportion of daily physical activity located in parks with playgrounds in new zealand children. *Preventive Medicine*, 50(5):235–240, 2010.

[13] Armanda Rodrigues, Carlos Damásio, and José Emanuel Cunha. Using GPS logs to identify agronomical activities. In *Connecting a Digital Europe Through Location and Place*, pages 105–121. Springer, 2014.

[14] Carlos N. Silla, Jr. and Alex A. Freitas. A survey of hierarchical classification across different application domains. *Data Mining and Knowledge Discovery*, 22(1-2):31–72, 2011.

[15] Aixin Sun and Ee-Peng Lim. Hierarchical text classification and evaluation. In *Data Mining, 2001. ICDM 2001, Proceedings IEEE International Conference on*, pages 521–528. IEEE, 2001.

[16] Philip J Troped, Marcelo S Oliveira, Charles E Matthews, Ellen K Cromley, Steven J Melly, and Bruce A Craig. Prediction of activity mode with global positioning system and accelerometer data. *Medicine and Science in Sports and Exercise*, 40(5):972–978, 2008.

[17] Shuangquan Wang, Canfeng Chen, and Jian Ma. Accelerometer based transportation mode recognition on mobile phones. In *Wearable Computing Systems (APWCS10)*, pages 44–46. IEEE, 2010.

[18] Yu Zheng, Like Liu, Longhao Wang, and Xing Xie. Learning transportation mode from raw GPS data for geographic applications on the web. In *Proc. of the 17th International Conference on World Wide Web*, WWW '08, pages 247–256. ACM, 2008.