

PREDICCIÓN DE ACTIVIDADES DE LA VIDA DIARIA EN ENTORNOS INTELIGENTES PARA PERSONAS CON MOVILIDAD REDUCIDA

A. Bertomeu-Motos, S. Ezquerro, Juan A. Barrios, Luis D. Lledó, Francisco J. Badesa, N. Garcia-Aracil
Universidad Miguel Hernández, Elche (Alicante) 03202,
abertomeu@umh.es

Resumen

Las Actividades de la Vida Diaria (AVD) son aquellas actividades orientadas al cuidado, productividad y ocio. Las tareas básicas de una persona, que consisten en el cuidado personal, se pueden ver afectadas por una enfermedad degenerativa o el sufrimiento de un accidente cerebrovascular. Por este motivo, el proyecto “Adaptive Multimodal Interfaces to Assist Disabled People in Daily Activities (AIDE)” pretende realizar un sistema multimodal capaz de extraer información relevante del comportamiento e intención del usuario, extraer información de las capacidades motoras residuales y del análisis del entorno con el objetivo de mejorar la independencia de estas personas. Para ello se ha realizado el estudio de dos técnicas de aprendizaje automático: campos aleatorios condicionales (CRF) y redes neuronales temporales (TDNN), para predecir la actividad que el usuario desea realizar estudiando también la adaptación de los métodos a la rutina del usuario. De este modo, y usando un entorno virtual, se ha adquirido la información necesaria de un usuario durante cinco días bajo las condiciones del proyecto AIDE obteniendo una tasa de acierto de más del 90 % en ambos métodos, llegando al 100 % de precisión en algunas actividades usando CRF. Por lo tanto, mediante esta técnica seremos capaces de ayudar a estas personas actuando sobre el hogar, adaptándonos a sus necesidades y rutinas.

Palabras clave: AVD, Movilidad Reducida, Entornos Inteligentes, Predicción, CRF, TDNN.

1. INTRODUCCIÓN

Las Actividades de la Vida Diaria (AVD) se definen como aquellas actividades orientadas al cuidado, productividad y ocio. Estas actividades se pueden dividir en dos grupos: las tareas básicas, consisten en el cuidado personal como pueden ser comer, beber o bañarse; y las tareas instrumentales, aquellas actividades que permiten una vida independiente en una comunidad como preparar comida, hacer la compra o el cuidado de personas [1].

El incremento de las técnicas de aprendizaje automático ha causado un gran interés en el desarrollo de entornos inteligentes de asistencia con funciones muy importantes como pueden ser la monitorización e intervención remota en ámbitos relacionados con el cuidado y la salud. En la literatura podemos encontrar estudios sobre el reconocimiento de actividades mediante entornos inteligentes, mediante técnicas de visión por computador a través de cámaras en primera persona o también realizando combinaciones entre cámaras y sensores de radiofrecuencia (RFID). Estas técnicas se centran en la ayuda de ancianos o personas con demencia [2][3][4][5].

Por otro lado, existen casas inteligentes con sensores embebidos capaces de generar información útil mientras un residente realiza su rutina diaria [6][7]. A partir de esta información, se pueden realizar técnicas de aprendizaje automático para predecir la actividad que el usuario está realizando aplicando modelos de aprendizaje que tengan en cuenta la información temporal de los datos. Estos modelos pueden ser modelos ocultos de Markov (HMM), campos aleatorios condicionales (CRF) o redes neuronales temporales (TDNN) entre otras [8][9].

En este trabajo se van a realizar técnicas de aprendizaje automático mediante dos modelos de predicción: CRF y TDNN. Estos modelos se estudiarán bajo las condiciones descritas en el proyecto “Adaptive Multimodal Interfaces to Assist Disabled People in Daily Activities (AIDE)” para realizar una predicción sobre la intención del usuario. Este proyecto pretende crear un sistema multimodal capaz de extraer información relevante de las capacidades motoras residuales, comportamiento e intención del usuario, del análisis del entorno y de factores de contexto para ayudar a personas con movilidad reducida que hayan sufrido algún daño cerebral o que tengan alguna enfermedad degenerativa. A través de una plataforma de simulación seremos capaces de obtener la localización del sujeto, los objetos con los que interactúa, la luminosidad de la estancia, la temperatura y la actividad que se está realizando necesarias para el entrenamiento de los modelos seleccionados.

Tabla 1: Tabla que relaciona las 10 actividades que se pueden realizar con los objetos a interactuar en cada una de ellas. Además, se muestran los objetos mínimos necesarios para realizar una correcta tarea.

ID	Actividad	Objeto	Objetos mínimos
1	Comer	Plato, tenedor, cuchillo, nevera, microondas	Plato y cuchillo
2	Beber	Vaso, botella, nevera	Vaso
3	Temperatura	Control de temperatura	Control de temperatura
4	Interacción con la TV	Televisión, mando a distancia	Mando a distancia
5	Entretenimiento	equipo de música, libros	Libros ó equipo de música
6	Persianas	Ventana	Ventana
7	Enchufar/apagar la luz	Interruptor	Interruptor
8	Lavarse los dientes	Cepillo de dientes, grifo	Cepillo de dientes
9	Lavarse la cara	Toalla, grifo	Toalla
10	Nada	-	-

2. ENTORNO VIRTUAL

Los datos necesarios para entrenar los modelos seleccionados se obtienen mediante una plataforma virtual capaz de simular la una rutina diaria controlada. Este entorno virtual se ha realizado dentro del proyecto AIDE, la Figura 1 muestra un frame del entorno virtual mientras se está realizando la tarea de comer. Como se puede observar, aparece un reloj que simula la hora del día en la que nos encontramos, avanzando a razón de un minuto por segundo. También realiza una detección automática de la localización del usuario dentro de la casa y del objeto que está observando. La actividad que se realiza se selecciona manualmente.

Además de esta información necesitamos conocer la luminosidad de la habitación y si el control de temperatura ha sido conectado o no. Esta informa-

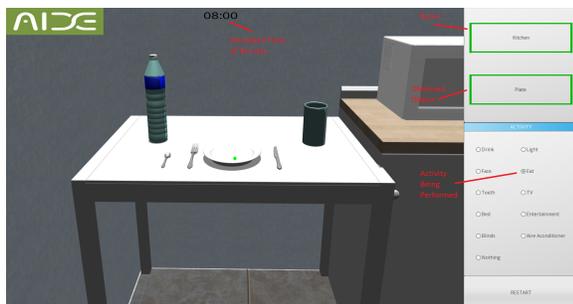


Figura 1: Frame del entorno virtual mientras se realiza la actividad de comer.

ción se obtiene a posteriori analizando los objetos con los que ha interactuado el usuario, obteniendo, de esta manera, la evolución de los dos parámetros a lo largo de la simulación. Las actividades, definidas en el proyecto, que se pueden realizar en cada estancia son

- Habitación: Enchufar/apagar la luz, levantarse, acostarse, subir/bajar persianas, interacción con la cama.
- Salón: Enchufar/apagar la luz, interacción con la televisión, interacción con el equipo de música, subir/bajar persianas, activar/desactivar temperatura.
- Baño: Enchufar/apagar la luz, lavarse la cara, lavarse los dientes.
- Cocina: Enchufar/apagar la luz, beber, comer, abrir/cerrar nevera, abrir/cerrar microondas.

En el entorno simulado se han simplificado las actividades para que resulte más fácil su uso simplificando el entorno virtual. Nos referiremos a las actividades ‘abrir/cerrar nevera’ y ‘abrir/cerrar microondas’ se asumen dentro de las actividades de *comer* y *beber* y se ha creado una actividad llamada *entretenimiento* a las actividades que se pueden realizar en el *salón* que no sean la interacción con la TV. Las actividades ‘levantarse’, ‘acostarse’, ‘interacción con la cama’ han sido eliminadas. En la Tabla 1 se reflejan los objetos mínimos con los que

Acquisition Time	Simulated Time	Room	Observed Object	Performed Task
20:45:04:	08:35	Kitchen	Table	Eat
20:45:05:	08:36	Kitchen	Table	Eat
20:45:06:	08:37	Kitchen	Plate	Eat
20:45:07:	08:38	Kitchen	Plate	Eat
20:45:08:	08:39	Kitchen	Plate	Eat
20:45:09:	08:40	Kitchen	Plate	Eat
20:45:10:	08:41	Kitchen	Plate	Eat
20:45:11:	08:42	Kitchen	Fork	Eat
20:45:12:	08:43	Kitchen	Fork	Eat
20:45:13:	08:44	Kitchen	Plate	Eat
20:45:14:	08:46	Kitchen	Plate	Eat
20:45:16:	08:47	Kitchen	Plate	Eat
20:45:17:	08:48	Kitchen	Knife	Eat
20:45:18:	08:49	Kitchen	Knife	Eat
20:45:19:	08:50	Kitchen	Table	Eat
20:45:20:	08:51	Kitchen	Table	Eat
20:45:21:	08:52	Kitchen	Bottle	Drink
20:45:22:	08:53	Kitchen	Bottle	Drink
20:45:23:	08:54	Kitchen	Bottle	Drink
20:45:24:	08:55	Kitchen	Bottle	Drink
20:45:25:	08:56	Kitchen	Bottle	Drink
20:45:26:	08:57	Kitchen	Wall	Drink
20:45:27:	08:58	Kitchen	Wall	Drink
20:45:28:	08:59	Kitchen	Glass	Drink
20:45:29:	09:00	Kitchen	Glass	Drink
20:45:30:	09:01	Kitchen	Glass	Drink
20:45:31:	09:02	Kitchen	Glass	Drink

Figura 2: Muestra de captura de datos.

se debe interactuar para realizar una actividad de manera correcta.

Mediante la Tabla 1 un usuario puede simular actividades que realizaría en su día a día y el sistema es capaz de almacenar esa información para después ser usada para entrenar el modelo de predicción deseado (ver Figura 2). Por lo tanto, este entorno de simulación es capaz de capturar cuatro variables de entrada: e, objeto, temperatura y luminosidad; y una variable de salida: actividad.

3. MODELOS DE PREDICCIÓN

Para el reconocimiento de las actividades que un usuario realiza y adaptarnos a su rutina diaria se necesitan modelos de aprendizaje automático que tenga en cuenta la información temporal de los datos que se han generado. Estos modelos se basan en el aprendizaje supervisado ya que, para un conjunto de datos de entrada, se conoce la clase a la que pertenece. Por lo tanto, se han elegido dos modelos de aprendizaje: CRF y TDNN.

Los CRF son un modelo probabilístico muy utilizado para segmentación y clasificación de datos[10]. Este modelo discriminativo tiene en cuenta no solo el estado actual sino que también utiliza los estados previos para realizar la predicción. Un modelo condicional calcula las probabilidades de una posible secuencia clasificada dada un conjunto de observaciones evitando así la restricción de independencia que asumen los modelos generativos como pueden ser los HMM[11].

Por otro lado, los TDNN son unas redes neuronales que trabajan sobre datos secuenciales. Esos operan, como otras redes neuronales, mediante múltiples capas de conexiones implementadas como una red neuronal *feedforward* en vez de como una red neuronal recurrente [12]. La principal ca-

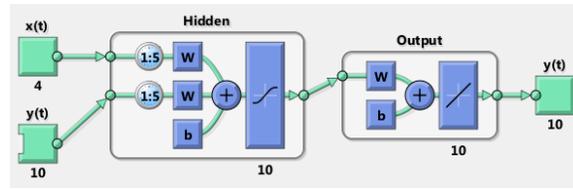


Figura 3: Estructura del TDNN con un tamaño de capa oculta de 10 y un retardo de 5 muestras.

racterística de este tipo de redes es la habilidad de expresar la relación entre las entradas de una secuencia temporal. Se van a usar las redes denominadas *Nonlinear Autoregressive with Exogenous input (NARX)*, una red de arquitectura dinámica [13]. La estructura del TDNN se puede observar en la Figura 3, donde se ha seleccionado un tamaño de capa oculta de 10 y un retardo de 5.

4. PRECISIÓN DE LOS MODELOS

Para probar los modelos seleccionados se ha usado el entorno virtual para generar datos. Estos datos representan a un usuario realizando actividades durante cinco días, 16 horas de actividad y 8 horas de descanso nocturno, intentando seguir la misma rutina todos los días. No había ningún requisito específico en el orden de interacción con los objetos a la hora de realizar una actividad específica. Por otro lado, los datos de entrada y salida de nuestro sistema se han clasificado en 5 estancias; un total de 24 objetos; temperatura activada o desactivada; luminosidad natural, artificial o sin luz; y un total de 10 tareas. En la Figura 4 se muestra la media de cada actividad con respecto a los cinco días y su desviación.

Como se pretende observar el comportamiento de los modelos con respecto a la rutina, se han realizado cuatro casos:

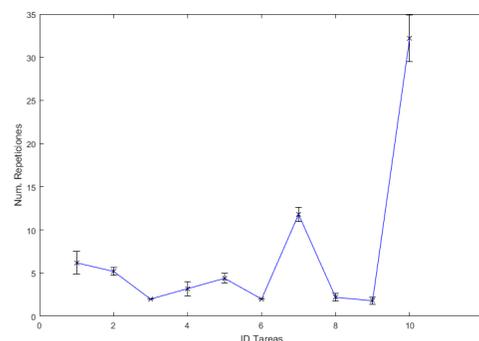


Figura 4: Media y desviación típica de las tareas con respecto a los 5 días.



(a) TDNN



(b) CRF

Figura 5: Matriz de confusión obtenida con los datos de test para los modelos en el caso 4vs1.

1. Un día para entrenamiento del modelo y un día para test del modelo (1vs1).
2. Dos días para entrenamiento del modelo y un día para test del modelo (2vs1).
3. Tres días para entrenamiento del modelo y un día para test del modelo (3vs1).
4. Cuatro días para entrenamiento del modelo y un día para test del modelo (4vs1).

La salida del TDNN se ha modificado de manera que sea un vector de dimensión 10 donde la salida correspondiente vale 1 y 0 en el resto de los casos. Los resultados de los distintos métodos implementados se muestra en la Tabla 2.

Se puede observar que el TDNN tiene una buena precisión mayor al 90% desde el primer día, sin embargo, el CRF es más susceptible a cambios en la realización de las tareas, variando su precisión a lo largo de los días. Sin embargo, el mayor porcentaje de acierto se obtiene en el caso 4vs1 mediante el CRF. En la Figura 5 se muestran dos matrices de confusión para el caso 4vs1 obtenida a partir de los datos de test para cada uno de los modelos.

Mediante la Figura 5 podemos ver fácilmente la

Tabla 2: Precisión de los modelos en cada caso.

Modelo	1vs1	2vs1	3vs1	4vs1
TDNN	92,6 %	93,1 %	93,1 %	92,5 %
CRF	84,5 %	92,3 %	86 %	93,9 %

debilidad de estos algoritmos. Podemos ver que las tareas de (1)comer y (1)beber, al ser muy parecidas, el TDNN comete un pequeño error mientras que el CRF no. Por otro lado, se puede observar que la tarea (7)Enchufar/apagar la luz se confunde fácilmente con la tarea (10)Nada en ambos casos, esto puede ser debido a que esta actividad puede ocurrir en cualquier estancia y en cualquier momento del día y, además, la única diferencia con respecto a la tarea (10) es la interacción con el interruptor. En ambos modelos, se puede observar que la predicción de las actividades que menos se realizan, como puede ser la que actividad (3)Temperatura, el TDNN tiene menor tasa de acierto que el CRF. Cabe añadir que, aunque la diferencia de precisión entre ambos métodos es pequeña, el CRF llega a tener un porcentaje de acierto del 100% en varias de las actividades mientras que el TDNN no.

5. CONCLUSIÓN

Este trabajo presenta una comparación de dos técnicas de aprendizaje automático con el objetivo de reconocer actividades de la vida diaria bajo las condiciones del proyecto AIDE. Los métodos que se han usado, CRF y TDNN, se basan en aprovechar la información temporal de los datos para obtener un mejor resultado. Se ha observado que ambos métodos tienen una gran tasa de acierto, por encima del 90% cuando se entrenan con cuatro días y se testea con un día. Sin embargo, el TDNN se adapta mejor a la rutina del usuario desde el primer día. Cabe destacar que,

aunque ambos modelos tengan una gran porcentaje de acierto, solamente el CRF llega a tener un 100 % de acierto en algunas actividades mientras que el TDNN solo alcanza un 97,4 % como máximo en la actividad (4) *Interacción con la TV*. Esto indica que el CRF es capaz de determinar con exactitud algunas actividades sin ninguna confusión y el TDNN tiene siempre un pequeño error de detección.

A partir de este trabajo podemos destacar que ambos métodos son muy buenos con respecto a la predicción de AVD en un entorno inteligente que sea capaz de extraer información sobre el comportamiento e intención del usuario, del análisis del entorno y de factores de contexto. De este modo, podremos ser capaces de ayudar a personas con movilidad reducida o con enfermedades degenerativas actuando sobre el hogar adaptándonos a las necesidades y rutinas de cada persona.

Agradecimientos

Este trabajo ha sido financiado por la Comisión Europea a través del proyecto AIDE: Adaptive Multimodal Interfaces to Assist Disabled People in Daily Activities (Grant agreement no: 645322) y por el Ministerio de Economía y Competitividad a través del proyecto DPI2015-70415-C2-2-R.

Referencias

- [1] M. Law, "Evaluating activities of daily living: directions for the future," *American Journal of Occupational Therapy*, vol. 47, no. 3, pp. 233–237, 1993.
- [2] H. Pirsiavash and D. Ramanan, "Detecting activities of daily living in first-person camera views," in *Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on*. IEEE, 2012, pp. 2847–2854.
- [3] S. Park and H. Kautz, "Hierarchical recognition of activities of daily living using multi-scale, multi-perspective vision and rfid," in *Intelligent Environments, 2008 IET 4th International Conference on*. IET, 2008, pp. 1–4.
- [4] M. Ghorbel, F. Arab, and M. Mokhtari, "Assistive housing: case study in a residence for elderly people," in *Pervasive Computing Technologies for Healthcare, 2008. PervasiveHealth 2008. Second International Conference on*. IEEE, 2008, pp. 140–143.
- [5] A. Lotfi, C. Langensiepen, S. M. Mahmoud, and M. J. Akhlaghinia, "Smart homes for the elderly dementia sufferers: identification and prediction of abnormal behaviour," *Journal of ambient intelligence and humanized computing*, vol. 3, no. 3, pp. 205–218, 2012.
- [6] D. J. Cook, A. S. Crandall, B. L. Thomas, and N. C. Krishnan, "Casas: A smart home in a box," *Computer*, vol. 46, no. 7, pp. 62–69, 2013.
- [7] T. L. Kasteren, G. Englebienne, and B. J. Kröse, "Human activity recognition from wireless sensor network data: Benchmark and software," *Activity recognition in pervasive intelligent environments*, pp. 165–186, 2011.
- [8] H. Alemdar, C. Tunca, and C. Ersoy, "Daily life behaviour monitoring for health assessment using machine learning: bridging the gap between domains," *Personal and Ubiquitous Computing*, vol. 19, no. 2, pp. 303–315, 2015.
- [9] S. T. M. Bourobou and Y. Yoo, "User activity recognition in smart homes using pattern clustering applied to temporal ann algorithm," *Sensors*, vol. 15, no. 5, pp. 11 953–11 971, 2015.
- [10] J. Lafferty, A. McCallum, F. Pereira *et al.*, "Conditional random fields: Probabilistic models for segmenting and labeling sequence data," in *Proceedings of the eighteenth international conference on machine learning, ICML*, vol. 1, 2001, pp. 282–289.
- [11] L. Rabiner and B. Juang, "An introduction to hidden markov models," *iee assp magazine*, vol. 3, no. 1, pp. 4–16, 1986.
- [12] V. Peddinti, D. Povey, and S. Khudanpur, "A time delay neural network architecture for efficient modeling of long temporal contexts." in *INTERSPEECH*, 2015, pp. 3214–3218.
- [13] J. M. P. Menezes and G. A. Barreto, "Long-term time series prediction with the narx network: An empirical evaluation," *Neurocomputing*, vol. 71, no. 16, pp. 3335–3343, 2008.