

Modelado de un sistema multi-agente aplicado a la predicción de la personalidad en Twitter

Christian Padilla-Navarro^{1,3}, Miguel Rebollo-Pedruelo¹, Carlos Lino-Ramírez²

¹ Universidad Politécnica de Valencia,
Departamento de Sistemas Informáticos y Computación, Valencia,
España

² División de Estudios de Posgrado e Investigación,
Instituto Tecnológico de León, León, Gto.,
México

³ Universidad Politécnica de Juventino Rosas,
Departamento de Telemática, Santa Cruz de Juventino Rosas, Gto.,
México

jopana2@doctor.upv.es, mrebollo@upv.es, carloslino@itleon.edu.mx

Resumen. En la presente investigación se realiza la simulación del comportamiento de un usuario de Twitter, a través de un sistema multi-agente, obteniendo las características de la personalidad de sus seguidores, aplicando el test de personalidad de los cinco grandes y los clasificadores Naive Bayes y KNN.

Palabras clave: Sistema multi-agente, test de los cinco grandes, Twitter, naive Bayes, KNN.

Multi-agent System Applied to Prediction of Personality on Twitter

Abstract. In this research the simulation of the behavior of a Twitter user is performed through a multi-agent system, obtaining the characteristics of the personality of his followers, applying the personality test of the Big Five and the classifiers Naive Bayes and KNN.

Keywords: Multi-agent system, big five test, Twitter, naive Bayes, KNN.

1. Introducción

Las redes sociales son hoy en día una de las formas de comunicación más importantes que hay en el mundo. En 2005 se estimaba que, sumando todas

las redes sociales existentes, había unos 115 millones de usuarios activos en el mundo [6], mientras que en 2015 se estimaba que tan solo Twitter tenía unos 316 millones de cuentas activas [5].

El vaciado de información que realizamos día con día nos enfrenta a expresar sentimientos, compartir opiniones, conversar con personas distantes, entre otros más, que nos ayudan a formar un “perfil social”. La red social Twitter, limitada a 140 caracteres por cada publicación (Tweet), es una de estas formas de expresión.

Durante años la psicología ha estudiado la personalidad del ser humano, y por tanto algunos investigadores han generado diferentes modelos para este fin. Las relaciones han sido descubiertas detrás de la personalidad y el desorden psicológico [13], el desempeño en el trabajo [1] y la satisfacción [9], e incluso algunos sucesos románticos [15].

Debido entonces a la gran cantidad de información vertida en las redes sociales, y a los millones de usuarios que pertenecen a ellas, el análisis de personalidad a través de redes sociales es uno de los objetos de estudio en tendencia. En [8] se aplicó el test de personalidad de los cinco grandes en la predicción de la personalidad de usuarios en Twitter. En [11] se aplicaron redes neuronales y el test de los cinco grandes aplicado a la predicción de la personalidad de usuarios de Twitter.

Simular el comportamiento de uno o varios usuarios en una red a través de agentes o de un sistema multi-agente ha sido un auxiliar importante en diversas investigaciones [18], [17].

Nosotros proponemos simular, a través de un sistema multi-agente, la interacción de un usuario de Twitter con sus seguidores aplicando clasificadores (KNN y Naive Bayes), utilizando las características del test de personalidad de los cinco grandes obtenidas de la API Watson de IBM Bluemix [2].

2. Antecedentes y trabajo relacionado

2.1. Test de personalidad de los Cinco Grandes (Big Five personality test)

Varios grupos de investigadores independientes descubrieron y definieron los cinco grandes factores mediante investigación empírica basada en datos. Ernest Tupes y Raymond Christal aportaron el modelo inicial, basado en el trabajo realizado en el Laboratorio de Personal de las Fuerzas Aéreas de los EE. UU. en la década de 1950 [16]. J. M. Digman propuso su modelo de los cinco factores de personalidad en 1990 [4], y Goldman lo llevó a los niveles más altos de las organizaciones en 1993 [7]. En un test de personalidad, para hacer referencia a los rasgos de los Cinco Grandes, también se puede utilizar el Modelo de los Cinco Grandes [3] y los Factores Globales de personalidad [12]. En la Figura 1 se muestran los cinco factores de personalidad.

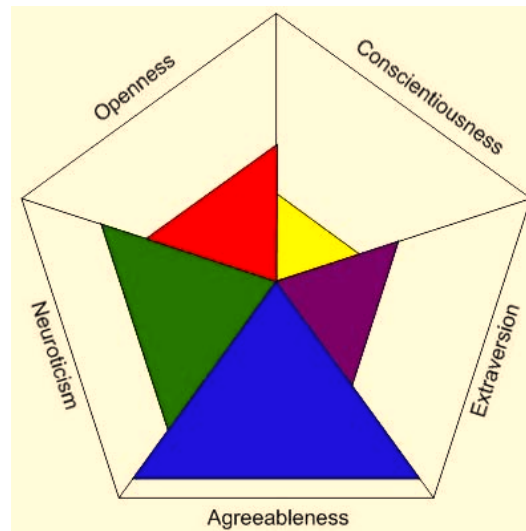


Fig. 1. Test de personalidad de los Cinco Grandes [8]

2.2. Rasgos del test de los Cinco Grandes factores de la personalidad

Apertura a nuevas experiencias. Apertura para apreciar el arte, emoción, aventura, ideas inusuales, imaginación, curiosidad, y variedad de experiencias. Las personas abiertas a las experiencias son intelectualmente curiosos, aprecian el arte, y son sensibles a la belleza. Comparados con personas cerradas tienden a ser más creativos y más concientes de sus sentimientos.

Responsabilidad. Es la tendencia a mantener autodisciplina, actuar responsablemente, y buscar cumplir los objetivos. Este rasgo denota preferencia por lo planeado antes que el comportamiento espontáneo. Influencia el modo en que controlamos, regulamos y dirigimos nuestros impulsos.

Extraversión. Esta caracterizado por emociones positivas y la tendencia a buscar estimulación en la compañía de otros. Un profundo compromiso con el mundo externo. Los extrovertidos disfrutan de la compañía de otros, y son percibidos como personas muy enérgicas. Tienden a ser entusiastas, individuos activos y en busca de emociones. En grupos, les gusta destacar, hablar y llamar la atención.

Amabilidad. Es la tendencia a ser compasivo y cooperativo en vez de antipático con los demás. La dimensión refleja diferencias personales en general, concernientes a la armonía en sociedad. Individuos amables valoran salir o juntarse con los

demás. Generalmente son considerados, amistosos, generosos, colaboradores y comprometidos con el deseo de los de su entorno.

Neuroticismo. Neuroticismo es la tendencia a experimentar emociones negativas como el enojo, la ansiedad o depresión. También es conocida como inestabilidad emocional.

2.3. Watson IBM BlueMix

El desarrollo en la nube IBM Watson ofrece una gran variedad de servicios para el desarrollo de aplicaciones cognitivas. Cada servicio de Watson proporciona una interfaz de programación de aplicaciones (API) para interactuar con este servicio.

IBM Bluemix [2] es la plataforma en la nube en la que se implementan las aplicaciones que se desarrollan en Watson.

Personality Insights. El servicio de percepción de personalidad de Watson de IBM (IBM Watson Personality Insights) proporciona una API que permite a las aplicaciones obtener puntos de vista de los medios sociales, datos de la empresa, u otras comunicaciones digitales. El servicio utiliza un análisis lingüístico para inferir la personalidad y características sociales, incluyendo los cinco grandes, necesidades y valores, a partir del texto. Estos conocimientos ayudan a las empresas a comprender las preferencias de sus clientes y mejorar su satisfacción mediante la previsión de sus necesidades, y además de recomendar acciones futuras. Las empresas pueden utilizar esta información para mejorar la captación de clientes, retención y compromiso, para fortalecer las relaciones con ellos [2].

3. Base de datos

Las características de cada usuario pueden ser vistas en la Tabla 1. Estas se obtuvieron a través del test de personalidad de IBM Bluemix y de la API de Twitter. Para generar dicha base de datos se analizó el perfil de 79 seguidores del usuario @jocripana, que es una cuenta real. El etiquetado de clases se realizó de forma manual por el usuario de la cuenta.

La base de datos utilizada para la clasificación de las características se encuentra formada de la siguiente manera:

3.1. Clasificación de la interacción

La clasificación recae directamente en cuatro clases: Clase 4, el agente tuitero debe seguir al agente seguidor y puede tuitear con él. Clase 3, el agente tuitero no debe seguir al agente seguidor, pero sí puede tuitear con él. Clase 2, el agente tuitero sigue al agente seguidor, pero no tuitea con él. Clase 1, el agente tuitero no sigue ni tuitea con el agente seguidor.

Tabla 1. Características para la clasificación de seguidores de Twitter

No.	Característica	Procedencia	Rango de Valor
1	Openness (Apertura)	Watson IBM Bluemix	0 a 100 %
2	Imagination (Imaginación)	Watson IBM Bluemix	0 a 100 %
3	Artistic Interests (Intereses artísticos)	Watson IBM Bluemix	0 a 100 %
4	Emotionality (Emocionalidad)	Watson IBM Bluemix	0 a 100 %
5	Adventurousness (Aventura)	Watson IBM Bluemix	0 a 100 %
6	Intellect (Intelecto)	Watson IBM Bluemix	0 a 100 %
7	Liberalism (Liberalismo)	Watson IBM Bluemix	0 a 100 %
8	Conscientiousness (Ser consciente)	Watson IBM Bluemix	0 a 100 %
9	Self-efficacy (Auto-eficacia)	Watson IBM Bluemix	0 a 100 %
10	Dutifulness (Sentido del deber)	Watson IBM Bluemix	0 a 100 %
11	Achievement-striving (Logro-esfuerzo)	Watson IBM Bluemix	0 a 100 %
12	Self-discipline (Auto-disciplina)	Watson IBM Bluemix	0 a 100 %
13	Cautiousness (Cautela)	Watson IBM Bluemix	0 a 100 %
14	Extraversion (Extroversión)	Watson IBM Bluemix	0 a 100 %
15	Friendliness (Amabilidad)	Watson IBM Bluemix	0 a 100 %
16	Gregariousness (Gregarismo)	Watson IBM Bluemix	0 a 100 %
17	Assertiveness (Asertividad)	Watson IBM Bluemix	0 a 100 %
18	Activity level (Nivel de actividad)	Watson IBM Bluemix	0 a 100 %
19	Excitement-seeking (Búsqueda de emociones)	Watson IBM Bluemix	0 a 100 %
20	Cheerfulness (Alegría)	Watson IBM Bluemix	0 a 100 %
21	Agreeableness (Afabilidad)	Watson IBM Bluemix	0 a 100 %
22	Trust (Confianza)	Watson IBM Bluemix	0 a 100 %
23	Morality (Moralidad)	Watson IBM Bluemix	0 a 100 %
24	Altruism (Altruismo)	Watson IBM Bluemix	0 a 100 %
25	Cooperation (Cooperación)	Watson IBM Bluemix	0 a 100 %
26	Modesty (Modestia)	Watson IBM Bluemix	0 a 100 %
27	Sympathy (Simpatía)	Watson IBM Bluemix	0 a 100 %
28	Neuroticism (Neuroticismo)	Watson IBM Bluemix	0 a 100 %
29	Anxiety (Ansiedad)	Watson IBM Bluemix	0 a 100 %
30	Anger (Enfado)	Watson IBM Bluemix	0 a 100 %
31	Depression (Depresión)	Watson IBM Bluemix	0 a 100 %
32	Self-consciousness (Auto-conciencia)	Watson IBM Bluemix	0 a 100 %
33	Immoderation (Falta de moderación)	Watson IBM Bluemix	0 a 100 %
34	Vulnerability (Vulnerabilidad)	Watson IBM Bluemix	0 a 100 %

Tabla 2. Clasificación de la interacción con el seguidor en Twitter

Clase	Lo sigo	Le escribo
Clase 1	X	X
Clase 2	X	✓
Clase 3	✓	X
Clase 4	✓	✓

4. Modelado del sistema multi-agente

4.1. Escenario 1

Para el primer escenario se definieron dos agentes dentro del sistema, el Agente Seguidor y el Agente Tuitero. Dichos agentes tienen atribuciones distintas.

Agentes seguidores. Los agentes seguidores, agentes artificiales, son generados de forma aleatoria. Pueden ser uno o varios, e intentan interactuar con el agente tuitero. Las tareas que tienen a cargo son las siguientes:

- **Enviar un Tweet al Agente tuitero.**- Sucede al ser generados. Es un evento que sucede de forma aleatoria.
- **Dar follow a al Agente tuitero.**- Sucede después de interactuar con el Agente tuitero. Es un evento aleatorio.
- **Responder a un Tweet.**- Sucede en caso de recibir un tweet del Agente tuitero.

Agente tuitero. El agente tuitero, agente artificial, es el centro de la toma de decisiones en el sistema. Este agente cubre con las siguientes tareas:

- **Seguir a un Agente.**- Sucede después de interactuar con el Agente seguidor, y sólo si después de la clasificación está permitida esta tarea.
- **Responder a un Tweet.**- Sucede en caso de recibir un tweet del Agente seguidor, y sólo si después de la clasificación está permitida esta tarea.

Modelado primer escenario. El modelado indica el proceso que seguirá el sistema multi-agente. Cuando llega un nuevo agente seguidor, envía un tweet al agente tuitero. El agente tuitero solicita a la API Watson de IBM Bluemix la extracción de las características del seguidor, a través del envío de un conjunto de tweets obtenido con la API de Twitter, aplicando el test de personalidad de los cinco grandes. Al tener las características las envía al clasificador Naive Bayes, que predice el tipo de interacción que el agente tuitero tendrá con el agente seguidor. Este modelado puede ser visto en la Figura 2.

4.2. Escenario 2

Para el segundo escenario se definieron tres agentes dentro del sistema, el Agente Tuitero, el Agente BlueMix y el Agente Clasificador. Dichos agentes tienen atribuciones distintas.

Agente tuitero. El agente tuitero, agente artificial, es el centro de la toma de decisiones en el sistema. Este agente cubre con las siguientes tareas:

- **Dar follow a un seguidor.**- Sucede después de interactuar con un seguidor, y sólo si después de la clasificación está permitida esta tarea.
- **Responder a un Tweet.**- Sucede en caso de recibir un tweet un seguidor, y sólo si después de la clasificación está permitida esta tarea.

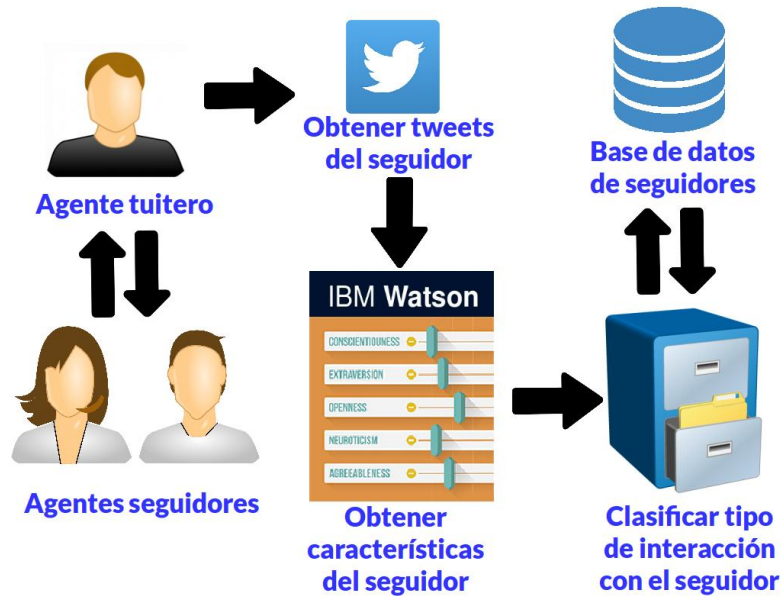


Fig. 2. Modelado del sistema multi-agente para el escenario 1

Agente BlueMix. El agente bluemix, agente artificial, se conecta con la api de Twitter para obtener una cantidad de tweets del seguidor con el que está interactuando. Posterior a esta acción, envía los tweets a la app de IBM Blue-Mix y extrae las características de los tweets del seguidor aplicando el test de personalidad de los cinco grandes (Watson).

Agente clasificador. Recibe del agente bluemix las características del seguidor con el que está interactuando el agente tuitero. Este agente tiene como tarea clasificar el comportamiento que debe tener el agente tuitero después de haber recibido un tweet de un seguidor o haber sido seguido por el mismo.

Modelado segundo escenario. Para el segundo escenario la interacción realizada es humano-agente. Llega un seguidor seguidor, envía un tweet al agente tuitero. El agente tuitero envía el usuario al agente bluemix. El agente bluemix solicita a la API Watson de IBM Bluemix la extracción de las características del seguidor, a través del envío de un conjunto de tweets obtenido con la API de Twitter, aplicando el test de personalidad de los cinco grandes. Al tener las características, las envía al agente clasificador, que este a su vez las manda al clasificador que realizará la predicción del tipo de interacción que el agente tuitero tendrá con el seguidor. Este modelado puede ser visto en la Figura 3.

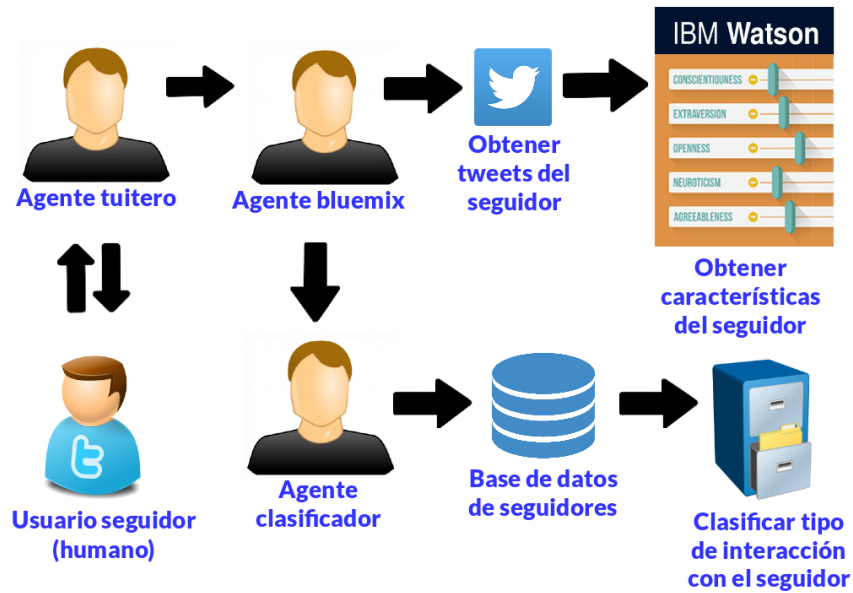


Fig. 3. Modelado del sistema multi-agente para el escenario 2

Tabla 3. Resultados del clasificador KNN

Vecinos	Capas	% de Clasificación
1	3	53.16 %
1	5	53.16 %
1	10	51.89 %
3	3	59.49 %
3	5	55.69 %
3	10	56.96 %
5	3	62.02 %
5	5	59.49 %
5	10	59.49 %
10	3	62.02 %
10	5	63.29 %
10	10	60.75 %

Tabla 4. Resultados del clasificador Naive Bayes

Capas	% de Clasificación
3	46.83 %
5	36.70 %
10	44.38 %

5. Experimentos y resultados

5.1. Clasificación

Para las pruebas previas de clasificación se utilizó el software Weka, y se aplicó en diversos clasificadores, buscando el mayor porcentaje de clasificación para su uso en posterior en la interacción en tiempo real. Las capas y los vecinos fueron propuestos de forma experimental.

6. Conclusiones y trabajo a futuro

Con la metodología propuesta se lograron obtener porcentajes de clasificación superiores al 50 % y en algunos casos incluso superiores al 60 % .

Para la simulación se utilizó SPADE, que es una plataforma libre, de sistemas multi-agente desarrollada en Python y basada en la tecnología de mensajería instantánea XMPP.

Como trabajo a futuro se buscará incluir más características, tales como la cadena de seguidores que tiene el usuario, así como propiedades de la red, el número de triángulos en la red, etc, con el fin de aumentar aún más el porcentaje de clasificación. De igual forma, se buscará aplicar un test similar para Facebook y el desarrollo de una plataforma propia de obtención de características de comportamiento.

Referencias

1. Barrick, M., Mount, M.: The Big Five personality dimensions and job performance: A meta-analysis. *Personnel psychology*, 44(1):1–26 (1991)
2. IBM. Información sobre la API Personality Insights de Watson de IBM Bluemix. <https://www.ibm.com/smarterplanet/us/en/ibmwatson/> (Consultado el 07 de Febrero del 2016)
3. Costa, P.T., Jr., McCrae, R.R.: Revised NEO Personality Inventory (NEO-PI-R) and NEO Five-Factor Inventory (NEO-FFI) manual. Odessa, FL: Psychological Assessment Resources (1992)
4. Digman, J.M.: Personality structure: Emergence of the five-factor model. *Annual Review of Psychology*, 41, 417–440 (1990)
5. El Mundo. Cuentas activas de Twitter en Septiembre del 2015. <http://www.elmundo.es/tecnologia/2015/09/24/5603bdcba474105398b4577.html> (Consultado el 07 de Febrero del 2016)
6. Golbeck, J.: Computing and Applying Trust in Web-based Social Networks. PhD thesis, University of Maryland, College Park, MD, USA (Abril 2005)
7. Goldberg, L.R.: The structure of phenotypic personality traits. *American Psychologist*, 48, 26–34 (1993)
8. Golbeck, J.: Predicting Personality from Twitter. In: IEEE International Conference on Privacy, Security, Risk, and Trust, and IEEE International Conference on Social Computing. DOI: 978-0-7695-4578-3/11 (2011)
9. Judge, T., Higgins, C., Thoresen, C., Barrick, M.: The big five personality traits, general mental ability, and career success across the life span. *Personnel psychology*, 52(3):621–652 (1999)

10. Ocean8. Test de personalidad de los cinco grandes. <http://ocean.na8.com.ar/blog/2012/11/29/que-es-ocean> (Consultado el 07 de Febrero del 2016)
11. Pundlik, M.: A Neural Network Approach to Personality Prediction based on the Big-Five Model. *International Journal of Innovative Research in Advanced Engineering (IJIRAE)*, Chapter 8, Volume 2 (Agosto 2015)
12. Russell, M.T., Karol, D.: 16PF Fifth Edition administrator's manual. Champaign, IL: Institute for Personality & Ability Testing (1994)
13. Saulsman, L., Page, A.: The five-factor model and personality disorder empirical literature: A meta-analytic review* 1. *Clinical Psychology Review*, 23(8):1055–1085 (2004)
14. Arana, J. Palanca, J.: SPADE User's Manual. <http://www.javierpalanca.com/spade/manual/> (Consultado el 07 de Febrero del 2016)
15. Shaver, P., Brennan, K.: Attachment styles and the 'Big Five' personality traits: Their connections with each other and with romantic relationship outcomes. *Personality and Social Psychology Bulletin*, 18(5):536 (1992)
16. Tupes, E.C., Christal, R.E.: Recurrent Personality Factors Based on Trait Ratings. Technical Report ASD-TR-61-97, Lackland Air Force Base, TX: Personnel Laboratory, Air Force Systems Command (1961)
17. Vázquez, A., Barrio, I., Vázquez-Salceda, J., Pujol, J.M., Sangüesa, R.: An agent-based Collaboratory. In: ACAI'01, Advanced Course on Artificial Intelligence, Prague (Julio 2001)
18. Watts, D.J., Strogatz, S.H.: Collective dynamics of 'small-world' networks. *Nature*, 393 (1998)