

Detección del nivel de confianza en las habilidades de un asistente virtual a partir del habla de la persona

Lara Gauder (0000-0001-6242-1546, mgauder@dc.uba.ar)^{1,2,*},
Leonardo Pepino (0000-0001-5037-3700, lpepino@dc.uba.ar)^{1,2},
Pablo Riera (0000-0001-8570-4688, priera@dc.uba.ar)¹,
Silvina Brussino (0000-0002-1087-644X, silvina.brussino@unc.edu.ar)^{3,4},
Jazmín Vidal (0009-0008-1306-7761, jvidal@dc.uba.ar)^{1,2},
Agustín Gravano (0000-0003-2812-6361, agravano@utdt.edu)^{5,6,7},
Luciana Ferrer (0000-0002-0426-8683, lferrer@dc.uba.ar)¹

¹ Instituto de Investigación en Ciencias de la Computación, CONICET-UBA, Argentina

² Departamento de Computación, Facultad de Ciencias Exactas y Naturales, UBA, Argentina

³ Facultad de Psicología, Universidad Nacional de Córdoba (UNC), Argentina

⁴ Instituto de Investigaciones Psicológicas, CONICET-UNC, Argentina

⁵ Laboratorio de Inteligencia Artificial, Universidad Torcuato Di Tella, Argentina

⁶ Escuela de Negocios, Universidad Torcuato Di Tella, Argentina

⁷ Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), Argentina

Resumen. Trabajos previos han demostrado que la confianza es un aspecto esencial en la interacción humano-computadora, determinando el grado en que una persona utilizará un sistema. La detección automática del nivel de confianza que un usuario tiene hacia un sistema podría permitir respuestas dinámicas para corregir situaciones de desconfianza. En este trabajo exploramos la factibilidad de detectar si el usuario está confiando en las habilidades de un asistente virtual (AV). Para ello, desarrollamos un protocolo novedoso para recolectar datos de habla de personas que interactúan con VAs con distintos niveles de habilidad. Los sujetos deben responder una serie de preguntas de conocimiento general mientras interactúan oralmente con un VA. El mismo es presentado junto con una calificación previamente asignada por otros usuarios como habilidoso o no y, además, responden de manera coherente a la puntuación reportada. El objetivo del protocolo es condicionar a los sujetos a un estado de confianza o desconfianza en las habilidades del VA, considerando que esto afectarán sus formas de habla que podrían ser detectados automáticamente. Recolectamos una base de datos de habla en español rioplatense que está disponible públicamente para fines de investigación. Los resultados obtenidos confirman que el protocolo logró inducir el estado mental deseado en los sujetos. Con los datos recolectados desarrollamos un sistema capaz de detectar la habilidad del VA con el que el sujeto interactuó durante una sesión utilizando características extraídas de las grabaciones de su habla con una precisión de más de 76%. Nuestro análisis sugiere que los sujetos modifican su forma de hablar según si confían o no en sus habilidades.

Palabras Claves: Confianza, Procesamiento del Habla, Interacción Humano-Computadora, Bases de Datos de Habla

DOI: <https://doi.org/10.1016/j.csl.2023.101487>

Publicado en: Computer Speech & Language, Volumen 60, páginas 101487, 2023

Factor de impacto: 3.1 (Q2)

Towards detecting the level of trust in the skills of a virtual assistant from the user's speech

Abstract. Research has shown that trust is an essential aspect of human-computer interaction directly determining the degree to which the person is willing to use a system. Automatic detection of the level of trust that a user has in a certain system could be used to enable adaptive responses to correct potential distrust. We aim to explore the feasibility of automatically detecting whether the user trusts the ability of the virtual assistant (VA). We developed a novel protocol for collecting speech data from subjects interacting with VAs with different skill levels. The subject is asked to respond to a series of factual questions while interacting orally with the VA. It is presented as having been previously rated by other users as either competent or incompetent and it answers the subjects' questions consistently to its ability. The goal of the protocol was to induce subjects to either trust or distrust the VA's skills, assuming that this would affect their speech patterns. We collected a speech corpus in Argentine Spanish which is publicly available for research use. Using the collected data, we developed a system to detect the ability of the VA with which a subject interacted during a session, based on the subject's speech patterns, with an accuracy up to 76%, compared to a random baseline of 50%. Our analysis suggests that subjects change the way they speak to the VA depending on whether they perceive it as more or less competent; that is, depending on whether they trust its ability.

Keywords: Trust, Speech processing, Human-computer interaction, Speech resources