



DOI: <http://dx.doi.org/10.23857/dc.v7i4.2167>

Ciencias técnicas y aplicadas

Artículo de revisión

Automatización de preselección en el área de talento humano, utilizando tecnologías de NLP y RPA

Automation of pre-selection in the area of human talent, using NLP and RPA technologies

Automação de pré-seleção na área de talento humano, utilizando tecnologias PNL e RPA

Jhonny Alexis Zapata-Chango ^I
jhonnyzapat@gmail.com
<https://orcid.org/0000-0002-7012-8291>

Pablo Marcel Recalde-Varela ^{II}
precalde@uisrael.edu.ec
<https://orcid.org/0000-0001-7256-2836>

Correspondencia: jhonnyzapat@gmail.com

***Recibido:** 30 de junio de 2021 ***Aceptado:** 15 de julio de 2021 * **Publicado:** 10 de agosto de 2021

- I. Tecnólogo en Análisis de Sistemas, Universidad Israel, Quito, Ecuador.
- II. Magister en Gestión de las Comunicaciones y Tecnologías de la información, MSC, Ingeniero de Sistemas de Computación E Informática, Universidad Israel, Quito, Ecuador.

Resumen

La presente investigación consiste en desarrollar un prototipo que resuelve la necesidad de clasificar y ordenar los currículos mediante la utilización de “Procesamiento de lenguaje natural” NLP y “Automatización robótica de procesos” RPA. El proceso consiste en extraer información no estructurada, consultar a fuentes de información dispersas y aplicar criterios de escala valorativa.

Para la extracción de la información no estructurada de los currículos se emplea, entrenamiento y creación de reglas con métodos tipo diccionarios.

Se despliega el modelo en el ambiente de IBM llamado “Watson Natural Language Understanding” NLU.

Para consultar las fuentes de información internas y externas del proceso, se emplea diagramación de bots identificados en el proceso de levantamiento de información.

Se construyen bots para consulta y consolidación de información acorde a las necesidades de escalas valorativas que son configurables por el “analista de recursos humanos” Selector.

Para el trabajo conjunto de bots, nlp y rpa se usa la tecnología de Rocketbot, IBM Watson Knowledge Studio, Python y 2Captcha.

Palabras claves: Curriculum; nlp; rpa; nlu; automatización; robótica.

Abstract

The present research consists of developing a prototype that solves the need to classify and order curriculums through the use of “Natural Language Processing” NLP and “Robotic Process Automation” RPA. The process consists of extracting unstructured information, consulting scattered sources of information and applying criteria of a rating scale.

For the extraction of unstructured information from the curricula, training and creation of rules with dictionary-type methods is used.

The model is deployed in the IBM environment called "Watson Natural Language Understanding" NLU.

To consult the internal and external sources of information of the process, a diagram of bots identified in the information gathering process is used. Bots are built to consult and consolidate information according to the needs of value scales that are configurable by the "human resources

analyst" Selector.

For the joint work of bots, nlp and rpa the technology of Rocketbot, IBM Watson Knowledge Studio, Python and 2Captcha are used.

Keywords: Curriculum; nlp; rpa; nlu; automation; robotics.

Resumo

A presente pesquisa consiste no desenvolvimento de um protótipo que resolve a necessidade de classificar e ordenar currículos por meio do uso de PNL de “Processamento de Linguagem Natural” e RPA de “Automação de Processos por Robótica”. O processo consiste em extrair informações não estruturadas, consultar fontes esparsas de informações e aplicar critérios de escala de rating.

Para a extração de informações não estruturadas dos currículos, é utilizado o treinamento e a criação de regras com métodos do tipo dicionário.

O modelo é implementado no ambiente IBM denominado "Watson Natural Language Understanding" NLU.

Para consultar as fontes internas e externas de informação do processo, é utilizado um diagrama dos bots identificados no processo de coleta de informações.

Os Bots são construídos para consultar e consolidar informações de acordo com as necessidades das escalas de valor configuráveis pelo Seletor “analista de recursos humanos”.

Para o trabalho conjunto de bots, nlp e rpa são utilizadas as tecnologias Rocketbot, IBM Watson Knowledge Studio, Python e 2Captcha.

Palavras-chave: Currículo; nlp; rpa; nlu; automação; robótico.

Introducción

Los procesos de reclutamiento y selección son parte fundamental en una organización, ya que permiten descubrir a los mejores talentos y construir equipos consistentes e innovadores.

El proceso se fundamenta en un grupo de procedimientos que buscan captar postulantes competentes y aptos para ocupar las diferentes plazas de trabajo publicadas por una organización.

En este proceso se encaminan actividades como aumentar el número de personas, tipo de personas que aplican a la vacante.

Automatización de preselección en el área de talento humano, utilizando tecnologías de NLP y RPA

Según (Universidad José Matías Delgado, 2016), las organizaciones toman decisiones en base a políticas de personal y fuentes de reclutamiento.

Políticas de personal: Se encuentran orientadas en influir la naturaleza de la vacante dentro de la organización.

Fuentes de reclutamiento: se encuentra basada en las fuentes de la organización donde se obtienen los candidatos innovadores y creativos utilizando fuentes internas y externas.

Para el caso de las políticas de personal se obedecen ciertas leyes laborales, preservando tanto a la imagen de la organización como al candidato, con el fin de evitar actos discriminatorios durante el proceso de selección.

Según (Mata, 2011), para expresar de forma correcta las necesidades de la empresa con respecto al puesto de trabajo, es ineludible conocer las fases que son parte del proceso de selección. Una de las etapas de selección, es la preselección de candidatos el cual supone revisar y analizar si los candidatos están acorde a las políticas de personal, además, si son idóneos al puesto de trabajo expuesto. Entre los requisitos, pueden existir requisitos indispensables o imprescindibles de modo que los no aptos, son inmediatamente eliminados del proceso, para esto, los selectores analizan cada curriculum vitae examinando mediante fuentes internas y externas de información para validar y puntuar cierta información descrita en el curriculum en la cual cada candidato obtendrá una calificación. Este proceso, al realizarlo en grandes cantidades, comprende el esfuerzo de varios selectores e inversión de tiempo.

Para un análisis ágil en el proceso de selección, hoy en día los avances tecnológicos proponen nuevas herramientas que apoyan a la lectura de texto no estructurado y automatizaciones que complementan la información de los perfiles postulantes. El “Procesamiento de lenguaje natural” NLP (Vajjala, 2020), es el permite extraer información no estructurada y la “Automatización robótica de procesos” RPA, permite automatizar los procesos que realiza un ser humano, aprovechándose de las herramientas habituales del analista.

Problema

Actualmente la etapa de preselección es muy importante para el proceso de selección de personal, ya que es el primer filtro que indica que postulantes están acorde al puesto ofrecido por la organización. Al existir mucha demanda de una plaza de trabajo, la persona encargada de realizar

Automatización de preselección en el área de talento humano, utilizando tecnologías de NLP y RPA

el análisis de curriculums (Selector), es propenso a cometer varios errores operativos al enfrentarse al consumo y procesamiento de varias fuentes de información. Esto se debe al desgaste provocado por la realización de tareas repetitivas.

(Matinez, 2016), afirma que: “El estancamiento o falta de progreso: es muy común encontrar a trabajadores que llevan años realizando la misma tarea”.

El desarrollo del prototipo mediante “Automatización Robótica de Procesos” RPA y “Procesamiento de lenguaje Natural” NLP no solo beneficiara a los Selectores implicados en el proceso, sino directamente al negocio al generar operaciones más eficientes, al ser una solución tecnológica ágil que reduce tiempos de ejecución, mayor precisión y ahorro de costos.

Objetivos y Justificación

Objetivos General

Automatizar la clasificación y ordenamiento de currículos orientados al área de tecnología del sector financiero, mediante la extracción de información no estructurada con “Procesamiento del lenguaje natural” NLP y la consulta de información en páginas públicas y sistemas de información interna por medio de “Automatización robótica de procesos” RPA.

Objetivos Específicos

- Desarrollar un modelo basado en reglas con la herramienta Watson Knowledge Studio para la extracción de datos personales, habilidades y herramientas manejadas por el postulante.
- Consolidar la información extraída de fuentes de información internas y externas, mediante bots construidos y orquestados por la herramienta Rocketbot.
- Aplicar criterios de escala valorativa para la clasificación de curriculums, en base a información recopilada y mediante tablas de asignación de pesos.

Justificación

Con el uso del procesamiento natural de lenguaje de NLP para la extracción de información no estructurada y, la aplicación de procesos automatizados tipo robots de RPA para la consulta masiva de fuentes de información pública se acelerarán los procesos de preselección y selección de curriculums idóneas para las vacantes a cubrirse.

Métodos

La automatización robótica de procesos (rpa) (Srivastava, 2021) es una forma de automatización que responde a las acciones del ser humano interactuando con la interfaz de usuario de un sistema principalmente software. Para la preselección de candidatos para una vacante los Selectores, deben revisar varias fuentes de información que contrastan con la información que los candidatos incluyen en sus curriculums, este proceso requiere de la lectura de lo entregado, el contraste con su registro profesional en la senescyt, la verificación de su pasado legal, la confirmación de que es o no una persona expuesta políticamente y que no constata con restricciones de nepotismo. Adicionalmente cada uno de estos hallazgos deben ser puntuados en base a una matriz que el Selector maneja y configura para un cargo o vacante a llenar.

En base a este proceso se diseñan seis bots: Uno llamado el padre que controla la interacción de los otros cinco.

El bot de procesamiento natural de lenguaje se encarga del manejar el proceso de seleccionar las palabras claves como los nombres, la cédula, el título académico, las experiencias laborales, entre otros. Para que este bot sea eficiente en tiempo de procesamiento se emplea NLU de IBM Watson, un servicio que para pruebas es gratuito y que para procesamiento masivo tiene costo.

El siguiente bot, es el de consultas a la SENESCYT, entidad ecuatoriana de educación superior que registra los títulos de nivel tecnológico y superior. Este bot encuentra los títulos y los almacena para que el candidato participe contrastando con la escala valorativa para una vacante.

Se diseña al bot de personas políticamente expuestas, revisa si el candidato es o fue parte de un cargo público a fin de valorar temas de idoneidad o riesgo principalmente político.

Se diseña al bot de nepotismo, el cual debe revisar si el candidato tiene alguna relación directa o de consanguinidad que puedan ser mal interpretadas al obtener dicha vacante y sus jefes organizacionales.

Finalmente se define al bot de antecedentes, el cual revisa los antecedentes penales de los candidatos para validar que no exista un historial que pueda afectar su postulación al cargo.

Cada bot interactúa con el bot padre quien organiza los resultados y lo contrasta con una matriz de valoración, misma que es definida por el Selector en base a las necesidades del cargo que requiere cubrir.

Automatización de preselección en el área de talento humano, utilizando tecnologías de NLP y RPA

Esa información es almacenada y se genera un reporte que es mostrado en orden de prelación de cada candidato que cumple con las condiciones para la vacante.

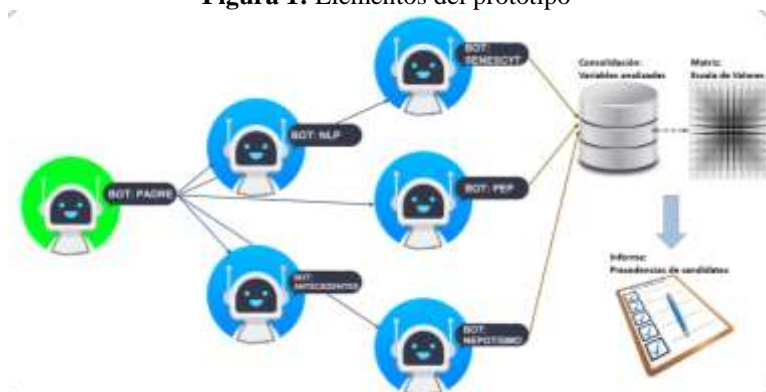
La población con la cual se hicieron las pruebas son las hojas de vida entregadas a analista de recursos humanos y consta de cien hojas de vida.

El estudio se lo realizó de forma experimental con los archivos de hojas de vida de una empresa real que requería de este proceso para cubrir sus vacantes.

Para desarrollar este prototipo se requiere de procesamiento natural de lenguaje (Yokota, 2019) que se lo desarrollo en la plataforma de IBM llamada IBM Watson Natural Language Understanding. Se emplea la tecnología de automatización robótica de procesos de Rocketbot. Y una gestión de almacenamiento con las bases de datos SQLite propia de Rocketbot para los procesos robóticos y Ms-SQL 2019 Server Community para almacenar el resultado de los procesos. Adicionalmente se utiliza Python 3.8.6 para controlar las variables que se procesan con cada bot y la estructura de archivos de procesamiento y logs. Finalmente se emplea un componente de gestión de captchas llamado 2Captcha que es un servicio SaaS, para vencer los validadores de cada página web que los tiene.

Finalmente, al contrastar todos los archivos procesados y comparados, se genera un informe con la precedencia de los candidatos que pueden cubrir un cargo para su posterior entrevista.

Figura 1: Elementos del prototipo



Nota: Orquestación de bots de rpa para la selección de candidatos, fuente propia.

Análisis de Resultados

El proceso de análisis básicamente es del tipo observación comparando los tiempos que le toman a un Selector humano procesar un grupo de candidatos en contraste a lo que le toma al prototipo realizar el mismo proceso.

Automatización de preselección en el área de talento humano, utilizando tecnologías de NLP y RPA

Comparando cien currículos a procesar por un solo Selector versus la misma cantidad con un solo bot padre, Tabla 1; se observa que en el caso del procesamiento humano tarda aproximadamente treinta horas o lo que equivale a dieciocho minutos por hoja de vida versus siete punto cinco horas equivalentes a cuatro puntos cinco minutos por curriculum con el bot.

Tabla 1: Tiempos de ejecución por curriculum

Interacción	Cantidad de Curriculums	Cantidad de selectores	Tiempo por 1 curriculum (minutos)	Tiempo total por curriculums (horas)
Humana	100	1	18	30,0
Bot	100	1	4,5	7,5

Nota: La información detallada es de fuente propia y es obtenida a partir de las pruebas realizadas.

Como se puede observar el procesamiento mediante bots, es cuatro veces más veloz que lo que un humano puede realizar.

Discusión de resultados

Este prototipo utilizar la versión Community de Rocketbot (Rocketbot RPA, 2021), misma que no cuenta con procesamiento paralelo. Considerando esta posibilidad se puede esperar que la velocidad de procesamiento no solo cuadruple la actual velocidad, sino que lo supere por un factor potencial de mínimo ocho veces lo actual, con lo cual el procesamiento del área de Selectores se convierta en un centro masivo de revisión de hojas de vida obteniendo a los mejores candidatos posibles.

Además, es importante mencionar que, al tener la información consolidada en una base de datos, abre a posibilidades de implementación de Machine Learning, permitiendo la mejora óptima del proceso de preselección e incluso, aportando valor a otras áreas de las instituciones.

Conclusiones

La automatización implementada en el proceso de preselección, fue construida de manera modular para el correcto mantenimiento del mismo, lo que permite que a futuro se puedan agregar nuevas funcionalidades haciendo esta solución escalable.

Automatización de preselección en el área de talento humano, utilizando tecnologías de NLP y RPA

Mediante la automatización se incrementó la capacidad de recepción de curriculums, ya que se reduce en un setenta y cinco por ciento el tiempo de ejecución de un solo recurso, tal como se demuestra en la sección de pruebas realizadas.

Al ser un Bot no atendido es posible parametrizar su ejecución en horarios no laborables, permitiendo que la información esté disponible cuando el recurso humano lo necesite, feriados, horas de la noche, y en épocas de máxima demanda de análisis de curriculums.

El beneficio más importante identificado en la implementación es el ahorro de costos en personal, ya que se mejora el tiempo invertido de 7.5 horas ejecutadas por los 4 recursos Selectores, a 6.7 horas ejecutadas con un solo Bot. Esto permite que el personal Selector enfoque su esfuerzo a tareas que generen valor.

Recomendaciones

Se recomienda agregar más fuentes de información tanto internas como externas que permitan enriquecer su base de conocimiento, y con esto generar modelos de Machine Learning que permitan aplicar un análisis óptimo.

Es importante que a futuro el bot tenga la capacidad de funcionar de manera atendida, con esto lograría que los Selectores puedan ejecutar casos que llegaran a destiempo.

Con la base de conocimiento generada por el bot, se recomienda generar modelos de Machine Learning que permitan mejorar el proceso de preselección, permitiendo la utilización de más parámetros que permitan la clasificación de curriculums de manera óptima.

Se recomienda que, para futuras fases, se implemente una gráfica que permita ver el estado de recolección de información de los curriculums y mediante dicha interfaz integrarse a los siguientes procesos que van más allá de la preselección. Esto permitirá tener una visión trescientos sesenta del proceso de Selección de personal.

Referencias

1. Mata, M. G. (2011). Importancia del reclutamiento y selección de personal en la toma de desiciones en las MPyMEs. Universidad Veracruzana, Facultad de Contaduría y Administración. Veracruz: Documento Interno.

Automatización de preselección en el área de talento humano, utilizando tecnologías de NLP y RPA

2. Matinez, M. A. (07 de 03 de 2016). iebschool. Obtenido de <https://www.iebschool.com/blog/desmotivacion-laboral-relaciones-laborales/>
3. Rocketbot RPA. (junio de 2021). Rockebot Studio. Obtenido de Construye tu robot de manera rápida y sencilla en MacOSX, Linux o Windows.: <https://www.rocketbot.com/es/>
4. Srivastava, V. (2021). Getting started with RPA using Automation Anywhere: Automate your day-to-day Business Processes using Automation Anywhere. New Delhi: BPB Publications.
5. Universidad José Matías Delgado . (01 de 04 de 2016). Gestión del telento humano como herramienta de competitividad en las empresas. Recuperado el junio de 2021, de <https://webquery.ujmd.edu.sv/siab/bvirtual/Fulltext/ADAG0000538/>
6. Vajjala, S. (2020). Practical Natural Language Processing: A Comprehensive Guide to Building Real-World NLP Systems. Sebastopol, CA.: O'Reilly.
7. Yokota, M. (2019). Natural Language Understanding and Cognitive Robotics. Boca Raton, FL: CRC Press.

©2020 por los autores. Este artículo es de acceso abierto y distribuido según los términos y condiciones de la licencia Creative Commons Atribución-NoComercial-CompartirIgual 4.0 Internacional (CC BY-NC-SA 4.0) (<https://creativecommons.org/licenses/by-nc-sa/4.0/>).