

Sistema de recomendación vocacional en línea

Sistema profissional recomendação on-line

Sergio Alcaraz Corona

Universidad Autónoma de Nuevo León, México

sergio.alcarazcrn@uanl.edu.mx

Mario Alejandro Bustos Jiménez

Universidad Autónoma de Nuevo León, México

mariobustosjmz@gmail.com

Resumen

Actualmente elegir la profesión deseada se ha convertido en una de las decisiones más difíciles dada la gran variedad de opciones e información disponibles. el proceso para seleccionar una profesión comienza con la correcta elección del programa educativo que más se ajuste a las habilidades, conocimiento y preferencias del individuo. para apoyar dicho proceso existen diferentes opciones tales como cursos de orientación vocacional así como aplicaciones informáticas, las cuales ayudan a los usuarios en su proceso de selección (katz, 1993); sin embargo, la mayoría de estas aplicaciones tienen un costo alto o no ofrecen resultados confiables. el objetivo de este trabajo es presentar un sistema de recomendación vocacional en línea basado en un algoritmo de asociación y clasificación de información como el principal componente de definición de perfiles vocacionales.

Palabras clave: sistemas de recomendación, orientación vocacional, inteligencia artificial.

Resumo

Actualmente escolher a profissão desejada tornou-se um dos dada a grande variedade de opções e informações de decisões mais difíceis disponíveis. o processo de seleção de uma profissão começa com a escolha certa de programa educacional que melhor se adapte as habilidades, conhecimentos e preferências do indivíduo. para apoiar este processo, existem diferentes opções, tais como cursos e aplicações profissionais, que ajudam os usuários em seu processo de seleção (Katz, 1993); No entanto, a maioria destas aplicações são caros e não fornecem resultados fiáveis. O objetivo deste trabalho é apresentar um sistema de recomendação on-line profissional com base em um algoritmo de associação e classificação da informação como o principal componente da definição de perfis profissionais.

Palavras-chave: sistemas de recomendação, aconselhamento profissional, inteligência artificial.

Fecha recepción: Julio 2015

Fecha aceptación: Noviembre 2015

Introdução

Uma das decisões mais importantes na vida de qualquer pessoa é selecionar sua profissão corretamente. A dificuldade em fazer uma decisão tão tem aumentado devido a vários fatores, principalmente econômicos, tecnológicos e sociais. Num mundo em mudança e cada vez mais se move mais rápido, a diversidade de opções de carreira tornou-se mais ampla e dinâmica, devido ao surgimento de novas profissões.

Recentemente, tem havido grandes mudanças nos currículos das diversas carreiras que oferecem quase todas as universidades, a fim de produzir profissionais mais bem preparados para um ambiente cada vez mais competitivo e globalizado. Ao mesmo tempo, novas carreiras foram surgindo para atender às novas demandas e exigências de trabalho. Isto, juntamente com outros

fatores, é um monte de informações que pode tornar o processo de seleção de uma carreira é uma tarefa difícil, especialmente se há elementos ou ferramentas de apoio.

Uma das opções para fora lá para apoiar o processo de seleção do ensino superior são cursos de formação profissional tradicionais, desenvolvidas principalmente por instituições de ensino públicas e privadas em diferentes formatos e conteúdos. Obviamente, um dos componentes essenciais de tais cursos de formação profissional é o questionário com perguntas destinadas a fornecer ao usuário com a informação que é útil para o seu processo de escolha de carreira. Hoje é muito comum encontrar estes questionários e outras informações através de um aplicativo ou encontrar informações na Internet. No entanto, a maioria desses tipos de questionários e guias são muito caros e não produzem resultados confiáveis (Franca, 2012). Por isso, é essencial ter ferramentas de apoio que prestam serviços de orientação profissional de novos e melhores (Katz, 1993). A ascensão da tecnologia da informação, especialmente tudo relacionado à internet, tem contribuído para o avanço e aperfeiçoamento de um grande número de aplicações na vida diária.

Atualmente, as chamadas aplicações web podem cobrir uma ampla gama de necessidades, que vão desde a compra de vários itens, fazer uma reserva ou consultar informações gerais, entre outras coisas. Em certos tipos de aplicações web, o resultado de uma consulta pode ser acompanhada por uma ou mais recomendações que fornecem o usuário com mais informações para tomar a melhor decisão. Por exemplo, a linha portal de compras Amazon.com recomenda certa quantidade similar aos itens pesquisados. Estas recomendações são obtidos por meio de algoritmos especiais usados em aplicações de inteligência artificial chamada e extração de dados, cujo processo a informação que irá proporcionar ao utilizador. Há uma grande variedade de algoritmos e metodologias para o efeito, no entanto, um tipo destes algoritmos procura associar e classificar as informações relacionadas umas com as outras de acordo com certos critérios ou regras que permitem o uso fácil e uso de tais informações.

O objetivo deste trabalho é projetar e implementar um sistema de recomendação profissional online utilizando linguagens de programação e tecnologias de informação recentes que permitem que o sistema seja capaz de produzir resultados confiáveis ao gerar as melhores recomendações de estudo para cada usuário a aplicação. Portanto, o sistema de recomendação proposta neste projeto é uma orientação website profissional para uso principalmente da comunidade estudantil da Faculdade de Engenharia Mecânica e Elétrica (FIME) da Universidade Autónoma de Nuevo

León (UANL), cuja URL identificador é <http://www.fime.vocacionaluanl.com.mx>. Para fazer isso, o sistema deve basear-se nas respostas que correspondem às perguntas relacionadas com as habilidades, aptidões, interesses e gostos que registra o usuário, de modo que todos eles são processados e classificados pela lógica do sistema e, assim, determinar em cada caso o perfil profissional mais adequado.

DESENVOLVIMENTO

Fundo

Como o próprio nome sugere, o objetivo principal de qualquer sistema de recomendação é fornecer algumas informações de interesse em resposta a um determinado pedido ou busca pelo utilizador ou utilizadores (Adomavicius, 2005). sistemas de recomendação são parte de uma extensa área de ciência da computação que é comumente conhecido como inteligência artificial, cujo principal objectivo é a gestão e administração de conhecimento e informação através de sistemas ou aplicações com algum nível de inteligência para resolver problemas e tomar decisões dentro de um determinado contexto (Truemper, 2004). Entre os sistemas inteligentes mais comuns são sistemas especialistas (Truemper, 2004), redes neurais (Kasabov, 1996), mineração de dados (Witten & Frank, 2005) e lógica fuzzy (Kasabov, 1996), entre outros.

De certa forma, sistemas de recomendação são consideradas como um tipo especial de sistema pericial. Além disso, há uma grande quantidade de aplicações que sistemas de recomendação têm ajudado a melhorar os processos de procura e selecção de informação. Uma das aplicações mais crescimento tem tido para os benefícios para os usuários e as organizações que implementam é o e-commerce, principalmente para a compra de produtos e serviços comerciais (Leavitt, 2006 ;. Linden et al, 2003).

Embora não seja tão extensa como no domínio comercial, encontramos algumas aplicações de sistemas de recomendação para ambientes acadêmicos ou de ensino (Aher e Wolf, 2013; Sobecki, 2014). Em particular, um destes trabalho apresenta uma aplicação deste tipo de sistema, que também usa uma implementação do algoritmo k Meios como parte de seu mecanismo para recomendar cursos para estudantes sob certos critérios ou condições especiais (aher e Lobo, 2013).

Uma das características distintivas dos sistemas de recomendação refere-se à lógica que segue o procedimento que é responsável por produzir as suas recomendações e do qual as seguintes categorias (Adomavicius, 2005) são derivados:

Recomendações com base no conteúdo. É a categoria que exige conhecimento das características relacionadas ao conteúdo ou descrição ou itens que foram anteriormente recomendado ao produzir novas recomendações para um determinado usuário ou grupo de usuários.

Recomendações baseadas em filtragem colaborativa. É outra categoria onde apenas consideradas as recomendações feitas anteriormente para os usuários com gostos, interesses e / ou características semelhantes para gerar novas recomendações para um determinado usuário ou grupo de usuários, ou seja, nestes casos, o conteúdo ou a descrição dos elementos não recomendado é relevante.

regimes híbridos são aqueles que combinam funcionalidade das duas categorias anteriores. Neste caso, a lógica do sistema proposto recomendação é o tipo da primeira categoria, uma vez que analisa o conteúdo; Especificamente, o valor da resposta selecionada para cada questão como parte do processo de geração de recomendações.

Definição de Perguntas

O processo de definição de perguntas começou com uma revisão da literatura para compreender os aspectos mais importantes no desenvolvimento de questões que correspondem a uma orientação profissional questionário. De acordo com a consulta, verificou-se que existem vários fatores que influenciam a definição de questões em tais questionários, tais como habilidades, conhecimentos, as preferências, personalidade e interesses de cada indivíduo, que são o mais importante. Em geral, os factores acima referidos têm sido extensivamente estudados para formar a chamada teoria da orientação profissional Holanda, que encontrada, entre outras coisas, que a personalidade e competências desempenham um papel importante em qualquer processo de selecção de ensino superior e eles estão se referindo ao que um indivíduo é capaz de fazer no futuro (Betz et al., 1989; Gati et al., 2006; Holland, 1997).

Por outro lado, uma pesquisa online (Alcaraz et al., 2015), que grupos de estudantes universitários foi enviado para tentar identificar os fatores que mais influenciaram no seu processo de seleção de uma carreira também foi criado, bem como várias situações que os

levaram a fazer mudanças de carreira. Com a implementação desta pesquisa uma amostra de cerca de 200, o que poderia obter dados valiosos para a definição do sistema e, assim, ter mais elementos para desenvolver uma ferramenta para orientação profissional mais eficiente e confiável foi obtido.

A partir da revisão da literatura e os resultados do inquérito, procedeu-se a criação de 50 perguntas para cada corrida gerenciado pelo sistema, que inclui três corridas da Faculdade de Engenharia Mecânica e Engenharia Elétrica da Universidad em sua versão inicial Nuevo Leon: Engenheiro de Software Technology (ITS), Engenheiro de Sistemas Manager (IAS), Engenheiro em Eletrônica e Comunicação (IEC). Depois de analisar os mais relevantes para o processo de definição de conceitos, determinou-se que o número de perguntas anterior pode representar as principais características e atributos de cada corrida de forma eficaz.

As perguntas são organizados dentro de arquivos do Excel, começando com o mais geral ao mais específico. Uma vez que as questões foram resolvidas, níveis (tiers) foram planejadas para eles. Os níveis incluem 0 a 3, com 3 sendo o mais representativo de cada nível de carreira, enquanto 0 é o nível mais básico ou geral (núcleo comum). O nível de distribuição aproximada de perguntas é proposto (nível, o número de perguntas) (0, 10), (1, 15), (2, 15), (3, 10).

A distribuição acima pode variar dependendo de vários fatores, entre os quais estão antigo, história, documentação e demanda por uma determinada carreira. No entanto, ao organizar as perguntas na forma acima é evitado ter de implementar um processo adicional para classificá-los em grupos de acordo com o grau ou nível de representação que eles têm para uma determinada carreira. Caso contrário, ele teria sido necessário para implementar os meios algoritmo k (K-meios), que é capaz de classificar e agrupar um conjunto de n observações, atribuindo a cada um dos possíveis grupos K (Duda e Hart, 1973). Portanto, o conteúdo de cada ficheiro Excel perguntas armazenados numa base de dados com base na ordem mencionada.

Estrutura do sistema de recomendação

Dentro da cada vez maior área de desenvolvimento de sistemas de informação, há uma variedade de opções, línguas e tecnologia para construir este tipo de aplicações web ou páginas de internet ferramentas dinâmicas. foi utilizado principalmente PHP (PHP Hypertext Preprocessor) é uma linguagem de programação utilizada para fins gerais para o desenvolvimento de aplicações web

com conteúdo dinâmico cliente-servidor (Beati, 2011) arquiteturas para este projeto. A linguagem PHP oferece diversas vantagens sobre outras linguagens de programação utilizadas para desenvolver aplicações sobre o tipo procurado neste projecto, ou seja, uma aplicação web dinâmica, com um banco de dados. Na Figura 1 o diagrama de componentes de software que compõem o sistema e são responsáveis pela sua operação é mostrado.

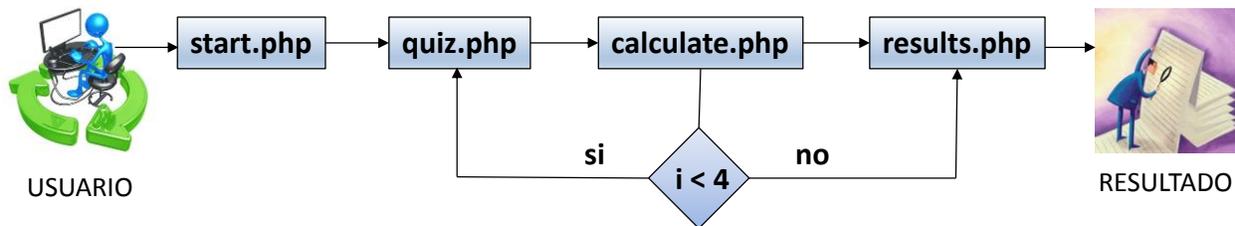


Figura 1. Estructura del sistema de recomendación fime.vocacionaluanl

O módulo start.php é responsável por estabelecer a conexão com o banco de dados e inicializado com variável de iteração zero (i) com a execução de módulos subsequentes são posteriormente controlada. Uma vez que a função acima, o módulo é executado quiz.php cuja principal função é exibir as perguntas da pesquisa começando com o nível mais geral (nível 0) para o usuário que você está respondida em 15 blocos por tela, como mostra a Figura 2. Note que as perguntas são todas de múltipla escolha e empregando uma escala de Likert (Likert, 1932) de 1 a 5, onde 1 representa o valor mínimo de associação ou de afinidade para uma determinada questão enquanto que 5 representa o valor máximo. Neste caso, a variável i representa o nível ao qual pertencem perguntas para ser exibido em qualquer dado momento.

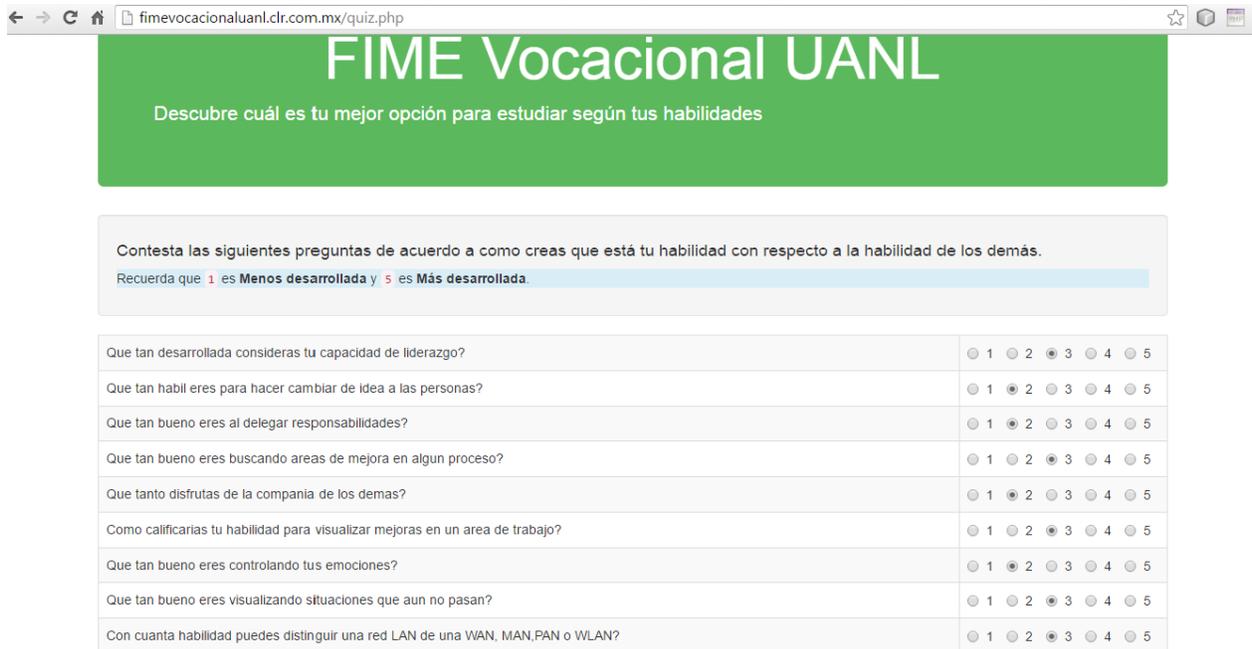


Figura 2. Funcionamiento del sistema de recomendación

Além disso, as perguntas que estão mostrando são extraídas a partir do sistema de base de dados, que consiste de duas tabelas, raças e perguntas. A Figura 3 mostra o diagrama da base de dados (fime.vocacionaluanlBD) utilizado pelo sistema. Por um lado, a raça tabela armazena dados para as corridas que suportam o sistema de partida com um identificador único para cada corrida, bem como o nome e data de cada corrida. Neste caso, a primeira raça (idcarrera = 1) no ser armazenados na base de dados foi IAS, seguido por IEC e finalmente STI. Por outro lado, a tabela de questão contém todas as declarações das perguntas e os seus respectivos níveis. Base de dados Fime.vocacionaluanlBD foi concebido e implementado com o MySQL, que é um banco de dados do sistema de gerenciamento de software livre ou open comumente usado no desenvolvimento de aplicações web, devido à sua funcionalidade, flexibilidade e confiabilidade (MySQL, 2016).

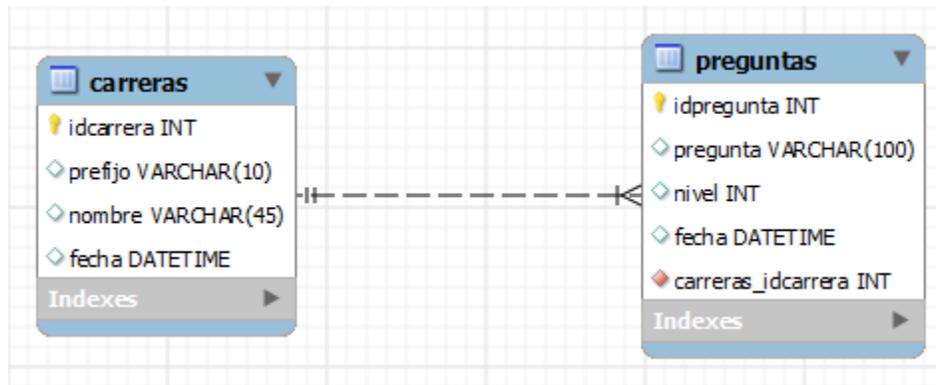


Figura 3. Esquema de la base de datos del sistema

Quando você terminar de responder todas as quinze perguntas em uma tela específica, quiz.php envia os valores definidos para estas perguntas para o módulo calculate.php a ser utilizado no cálculo dos pesos necessários e determinar o perfil profissional mais adequado o usuário. A Figura 1 mostra que, se a iteração (i) for inferior a 4, em seguida, o controle retorna ao quiz.php para exibir o próximo conjunto de perguntas em uma nova tela, mas não antes de aumentar o valor da variável de iteração i para e todos vêm de um nível superior. Se iteração não é inferior a 4, em seguida, o último módulo results.php é executado para exibir os resultados da avaliação, o que será explicado em pormenor na secção seguinte.

Os pesos acima mencionadas são necessárias para determinar o percentual de participação de cada questão a uma raça específica, ou seja, um total de 100% para cada associação de interrogação percentagens ou associação são obtidos para cada corrida. Como mencionado, a versão inicial do sistema processa três corridas, então deixe $m = 3$, onde m representa o número de etapas que suporta o sistema e n representa o nível. Por exemplo, para cada pergunta percentagens para IAS, IEC e STI carreiras, ou seja, o cálculo dos pesos é possível determinar se o IAS, IEC ou DSTs são uma questão foram alocados.

Tabela I mostra o procedimento para o cálculo dos pesos de acordo com o nível eo número de corridas suportados e pesos de capital para $n = 0$ porque o nível de questões mais gerais, que se aplica principalmente para todo o sistema de processamento de corridas. Em vez disso, a partir do primeiro nível em diante deve ver maior peso para a raça a que uma questão pertence. Por exemplo, para $n = 1$, o valor do coeficiente de correção $1 / (3 + 1) = 0,25$ é atribuído a cada uma das etapas que não têm relação com a causa e o valor restante para completar 1 é atribuído à raça

associado finalmente deixando o IAS = 0,25 distribuição, IEC = 0,25 e DST = 0,5 para uma pergunta STI carreira. Em geral, os pesos mostrados na Tabela I correspondem a questões de ITS para cada um dos diferentes níveis, de modo que os valores mais altos encontrados nesta coluna. Uma grande vantagem é que o sistema foi implementado de modo que seja capaz de se adaptar a qualquer número de corridas de forma dinâmica, que por sua vez facilita o cálculo dos novos pesos sem fazer alterações para a lógica do sistema cada vez que você precisa adicionar uma nova carreira e seu respectivo conjunto de perguntas.

Tabla I. Cálculo de ponderaciones para cada pregunta en base al nivel y al número de carreras

Nivel (n)	Ponderación	IAS	IEC	ITS
0	$1/(m + n)$	0.33	0.33	0.33
1	$1/(m + n)$	0.25	0.25	0.5
2	$1/(m + n)$	0.2	0.2	0.6
3	$1/(m + n)$	0.16	0.16	0.67

Resultados e discussão

Depois de apresentar uma visão geral do sistema de recomendação fime.vocacionaluanl profissional, nesta seção vamos nos concentrar no final da operação, quais são os resultados e as análises realizadas para validar o desempenho ea funcionalidade do sistema. Anteriormente, na Figura 4 que mostra o diagrama de sequência que se segue o sistema a ser usado para visualizar os diferentes fluxos de funcionamento normalmente realizada entre o utilizador, o sistema e o banco de dados. Como você pode ver, existem três principais fluxos de operação: tela, e pesar responder a perguntas, que são repetidos quatro vezes porque são quatro níveis de perguntas no total (0 a 3). O tempo médio que leva para responder a todo o questionário (60 perguntas) é cerca de 10 minutos.

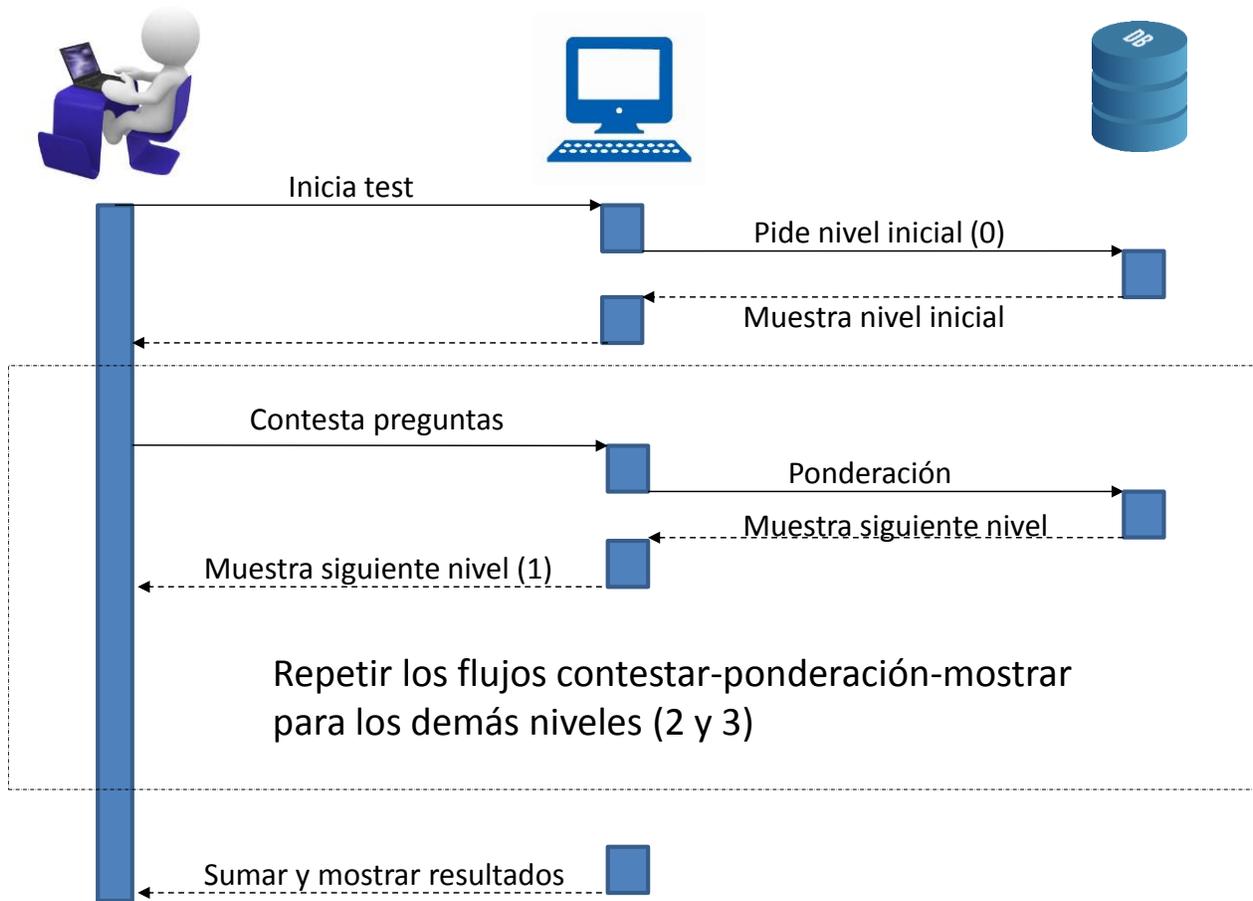


Figura 4. Diagrama de secuencia de operación del sistema fime.vocacionaluanl

O sistema está disponível para o momento em um espaço que foi contratado em um servidor para hospedar tanto o software do sistema e seu banco de dados. Além disso, o tipo de serviço que iria adaptar o sistema contratado recomendação de outras faculdades usando o nome da autoridade, seguido do nome de domínio vocacionaluanl para formar o seu URL. Para o design gráfico do sistema, decidimos usar Bootstrap que é um tipo especial de biblioteca (framework) de software livre projetado para facilitar a criação de páginas web dinâmicas. Com base em linguagens HTML e CSS, Bootstrap contém uma variedade de recursos para criar aplicações web interativas muito completo.

A Figura 5 mostra um exemplo de como os resultados parecem responder terminar todo o questionário. Neste caso, o nome da raça que determina o sistema mais compatível com as habilidades, conhecimentos e mais desenvolvido nos interesses do usuário aparece.



Figura 5. Pantalla de resultados del sistema

Finalmente, é importante concluir que apresenta uma análise mostrando que de alguma forma o sistema tem um desempenho correcto e consistente. Portanto, uma vez que o questionário representa a parte central da funcionalidade da recomendação do sistema, para testar a validade e eficácia da análise com o índice ou coeficiente alfa de Cronbach para uma amostra de usuários foi obtida entrevistados foi realizada o questionário. Este índice é comumente usado para determinar a consistência e confiabilidade de um instrumento de medição constituído por vários elementos, variáveis ou itens correlacionados (Cronbach, 1951). O intervalo de valores que pode assumir este coeficiente é entre 0 e 1, onde 0 representa ausência de uma confiabilidade enquanto confiabilidade absoluta. Tipicamente, um valor igual a ou maior do que 0,8 é considerada alta fiabilidade.

A amostra analisada ascendeu a 40 e foi tirada de uma população de estudantes que se formam STI e IAS, principalmente, que foram convidados a usar o sistema para determinar a sua fiabilidade. No entanto, é claro que o objectivo do sistema é, por um lado, orientar os alunos que se encontram ainda incerto sobre a corrida eles vão estudar e, por outro, o sistema é capaz de cobrir todas as carreiras de engenharia É oferecido pela faculdade em versões posteriores.

Das duas formas disponíveis para estimar a escala Cronbach Alfa, escolhemos o seguinte:

Conclusão

Neste artigo apresentamos uma recomendação on-line sistema profissional baseado em um algoritmo chamado de inteligência artificial como o principal componente para determinar o melhor perfil profissional de cada usuário. O sistema visa tornar-se uma ferramenta valiosa de apoio à livre utilização para todos aqueles estudantes que estão considerando a escolha de uma carreira de engenharia na FIME UANL. O sistema foi concebido e implementado com linguagens de programação e ferramentas amplamente utilizadas neste tipo de aplicações dinâmicas e interativas. Poderia ser ainda validar a consistência ea confiabilidade do sistema por meio de uma análise estatística com dados obtidos a partir de uma amostra de 40 usuários. Recomenda-se a aumentar no futuro o número de corridas FIME que podem apoiar o sistema, e até mesmo replicar o sistema para que este processo é executado outras faculdades da UANL. Isso tem a ver com a investigação e definir o conjunto de questões que correspondem a todas as raças que você deseja incluir e organizá-los sobre os níveis definidos de modo que tudo é armazenado no banco de dados.

Agradecimentos

Queremos agradecer os estudantes do ITS carreira UANL, Xavier Sanchez Diaz, Alfredo González Garza e Cynthia Reyna Elizabeth Villagomez, por seu envolvimento na pesquisa e desenvolvimento da primeira versão do sistema de recomendação profissional; Agradecemos também o apoio do Programa de Desenvolvimento Profissional para Professores (PRODEP) do Ministério da Educação.

Bibliografía

- Adomavicius, G., y Tuzhilin, A. (2005). Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions. *IEEE Transactions on Knowledge and Data Engineering*, 17(6), pp. 734-749.
- Aher, S. B., & Lobo, L. M. R. J. (2013). Combination of machine learning algorithms for recommendation of courses in E-Learning system based on historical data. *Knowledge Based Systems*, 51, pp. 1-14.
- Alcaraz, S., Sánchez, X., González, A. (2015). Encuesta orientauanl. FIME UANL. Sitio web: <http://bit.ly/orientaUANL>
- Beati, H. (2011). *PHP: Creación de páginas Web dinámicas*. Buenos Aires: Alfaomega Grupo Editor Argentino.
- Betz, N. E., Fitzgerald, L. F., Hill, R. E. (1989). Trait-factor theories: Traditional cornerstone of career theory. In M. B. Arthur, D. T. Hall, & B. S. Lawrence (Eds.), *Handbook of career theory*. Cambridge: Cambridge University Press, pp. 26-40.
- Cronbach, L. J. (1951). Coefficient alpha and the internal structure of test. *Psychometrika*, 16(3), pp. 297-334.
- Duda, R. O., & Hart, P. E. (1973). *Pattern classification and scene analysis*. John Wiley & Sons, Inc.
- Franca, D. (2012). Decide tu carrera, orientación vocacional para jóvenes mexicanos. Relpe. Sitio web: <http://www.relpe.org/decide-tu-carrera-orientacion-vocacional-para-jovenes-mexicanos/>
- Gati, I., Fishman-Nadav, Y., Shiloh, S. (2006). The relations between preferences for using abilities, self-estimated abilities, and measured abilities among career counseling clients. *Journal of Vocational Behavior*, 68, pp. 24-38.
- Holland, J. L. (1997). *Making vocational choices: A theory of vocational personalities Uand work environments*. Odessa, FL: Psychological Assessment Resources.

- Kasabov, N. K. (1996). Foundations of neural networks, fuzzy systems, and knowledge engineering. Cambridge: MIT Press.
- Katz, M. R. (1993). Computer-assisted career decision making. Hillsdale, NJ: Erlbaum.
- Leavitt, N. (2006). Recommendation Technology: Will It Boost E-Commerce? IEEE Computer Society.
- Likert, R. (1932). A technique for the measurement of attitudes. Archives of Psychology, 140, pp. 1-55.
- Linden, G., Smith, B., York, J. (2003). Amazon.com Recommendations Item-to-Item Collaborative Filtering, IEEE Internet Computing.
- Sobecki, J. (2014). Comparison of selected swarm intelligence algorithms in student courses recommendation application. International Journal of Software Engineering and Knowledge Engineering, 24(1), pp. 91-109.
- Truemper, K. (2004). Design of logic-based intelligent systems. Wiley-Interscience.
- Witten, I. H., & Frank, E. (2005). Data mining: Practical machine learning tools and techniques. Second Edition, Morgan Kaufmann.