Sistema para la Búsqueda y Clasificación de Noticias

***Trabajo Terminal No. — — — — -— — —***

*Alumnos: Hernandez Gomez Carlos Andres, Meza Martínez Luis Daniel*

*Directores: Juárez Gambino Joel Omar, García Mendoza Consuelo Varinia*

*Turno para la presentación del TT: Matutino*

*e-mail: ldanielmezam@gmail.mx*

### Resumen – Hoy en día existen distintas técnicas para la recuperación de información en la Internet, sin embargo, hay ocasiones en las cuales requerimos hacer una búsqueda utilizando palabras clave para identificar un tema, fecha o lugar en específico lo cual implica un tiempo mayor a la búsqueda. En el presente trabajo terminal se propone la realización de un sistema para la búsqueda y clasificación de noticias a través de una serie de filtros utilizando aprendizaje automático.

**Palabras clave** – Aprendizaje automático, búsqueda de texto, lenguaje natural, procesamiento de texto.

**1. Introducción**

Lenguaje natural[1]

Sistemas similares que se han desarrollado son.

1. Aplicación.
2. Proyecto de Investigación CGPI 2004-1276.
3. Proyecto académico realizado en otra institución.
4. TT.

Es recomendable añadir una tabla comparativa que enfatice las aportaciones del Trabajo Terminal que se propone. Se muestra un ejemplo en la Tabla 1.

|  |  |  |
| --- | --- | --- |
| **SOFTWARE** | **CARACTERÍSTICAS** | **PRECIO EN EL MERCADO** |
| Flipboard |  |  |
| Proyecto |  |  |
| TT |  |  |
| Solución Propuesta |  |  |

**Tabla 1**. Resumen de productos similares.

A continuación se mencionan algunas recomendaciones para la redacción de este documento. Se sugiere una redacción clara y precisa donde suele emplearse verbos en primera persona del plural: “hemos observado”, “ya decíamos que”, etc.; así como los verbos en formas impersonales: “se dice”, se produce”, “se observa”. Construcción lógica de oraciones: sujeto + verbo + complementos, evitando trastrocamiento del orden lógico, es decir, abordando cada apartado o sección para que vaya teniendo coherencia. Se emplean construcciones simples y párrafos breves. Las referencias se colocarán entre corchetes y se numerarán de acuerdo a su orden de aparición en el texto. Por ejemplo, para conocer más recomendaciones en la redacción del texto, se puede consultar el trabajo de Baena [1], y otros similares [2,3].

Se recomienda no exceder un máximo de cinco páginas, sin contar los cronogramas. Estimando para las secciones de la 1 a la 5 un total de cuatro páginas, se harán recomendaciones respecto a la extensión de las diferentes secciones. La introducción no debe exceder un 30% de dicho total, aproximadamente 1.2 páginas.

### 2. Objetivo

Consiste en un enunciado claro y preciso de lo que se pretende realizar, es decir, constituye la meta a lograrse. Indica, además, los alcances del trabajo. No debe exceder un 5% del total, aproximadamente 0.2 páginas.

**3. Justificación**

Se realiza el planteamiento del problema y se determina una solución tentativa a éste.

Describe las contribuciones o beneficios que resultarían con el TT (originalidad del trabajo, vinculación con los usuarios potenciales, utilidad de los resultados, complejidad en su elaboración, mejora a lo ya existente, etc.). Además, determina la viabilidad del TT en cuanto a: tiempos, recursos, alcances, costos, etc.

*Originalidad del trabajo*. Que se aporta de novedoso, distinto o recién propuesto en el TT.

*Vinculación* con los usuarios potenciales. Quiénes son directamente los beneficiados: un sector de la sociedad, una organización, el sector productivo, la comunidad “X”; es decir, para quiénes tendría interés.

*Mejora a lo ya existente*. De qué manera se va a intervenir para propiciar la mejora de algo que ya existe.

*Complejidad*. Que el TT refleje calidad en el desempeño y aplicación de los conocimientos adquiridos o acumulados a lo largo de la carrera, así como de las horas mínimas destinadas al trabajo.

Esta sección no debe exceder un 25% del total, aproximadamente una página.

**4. Productos o Resultados esperados**

En este apartado se describirán, de manera esquemática, los productos o resultados finales del TT, considerando el sistema a desarrollar, los documentos a generarse, las pruebas o experimentos a realizar para validar los resultados, etc.

Se recomienda esquematizar a grandes rasgos (diagrama de bloques) la arquitectura del producto a desarrollar en la Figura X.

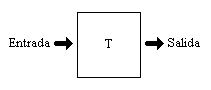


Figura X. Arquitectura del sistema.

También se pueden enlistar los productos esperados del TT, por ejemplo:

1. El código.
2. La documentación técnica del sistema.
3. El manual de usuario.
4. Alguna publicación.

Esta sección no debe exceder un 20% del total, aproximadamente 0.8 páginas.

**5. Metodología**

Se describirán las vías que facilitarán el desarrollo del proyecto. Se sugiere definir las metas intermedias a alcanzar y seguir algún estándar (por ejemplo, ISO9001), pero no se pondrá la descripción del estándar, sino su aplicación a la planeación del trabajo a realizar. Se mencionarán las técnicas y herramientas a emplear.

Esta sección no debe exceder un 20% del total, aproximadamente 0.8 páginas.

**6. Cronograma**

Planear las etapas, actividades o tareas clave que demande el TT y estimar su tiempo de realización. Su elaboración debe apoyar la delimitación de los alcances del proyecto. Elaborar un cronograma por alumno. El tiempo considerado debe concluir normalmente en Mayo, los casos especiales se revisarán de manera individual. Al final de este documento se encuentra un ejemplo.

**7. Referencias**

1. G. Baena Paz, S. Montero Olivares, Tesis en 30 días, Cd. Neza, EDOMEX, México, 1999.
2. S. Mercado, ¿Cómo hacer una Tesis?, México, D.F., México, 1994.
3. M. Tamayo y Tamayo, Metodología Formal de la Investigación Científica, México, D.F., México, 1997.
4. Título del Trabajo Termina, TT #####, Año de realización, Nombre del director.
5. Sistema para administración de información de centros comerciales, TT 0353-6, 2002, Sánchez Juárez J.
6. Título del Proyecto de Investigación o propuesta de estudio con registro ante CGPI, Nombre del director del proyecto, Proyecto de Investigación CGPI 20##-####.
7. INTRA-ESCOM Intranet para la Escuela Superior de Cómputo del Instituto Politécnico Nacional, Carreto Arellano C., Proyecto de Investigación CGPI 2002-0876.
8. Proyecto CONACYT #####, Nombre del responsable.
9. G. O. Young, “Synthetic structure of industrial plastics (Estilo de libro con título de la contribución y editor),” in Plastics, 2nd ed. vol. 3, J. Peters, Ed. New York: McGraw-Hill, 1964, pp. 15–64.
10. W.-K. Chen, Linear Networks and Systems (Estilo para libros). Belmont, CA: Wadsworth, 1993, pp. 123–135.
11. H. Poor, An Introduction to Signal Detection and Estimation. New York: Springer-Verlag, 1985, cap. 4.
12. B. Smith, “An approach to graphs of linear forms (Estilo para trabajos sin publicar),” sin publicar.
13. E. H. Miller, “A note on reflector arrays (Estilo para revistas especializadas—Aceptado para su publicación),” IEEE Trans. Antennas Propagat., en proceso de publicación.
14. J. Wang, “Fundamentals of erbium-doped fiber amplifiers arrays (Estilo para revistas especializadas—Enviado para su publicación),” IEEE J. Quantum Electron., enviado para su publicación.
15. F. A. Sánchez Garfias, Escuela Superior de Cómputo, Instituto Politécnico Nacional, México, D.F., México, comunicación personal, Mayo 1995.
16. Y. Yorozu, M. Hirano, K. Oka, and Y. Tagawa, “Electron spectroscopy studies on magneto-optical media and plastic substrate interfaces (Estilo para traducciones de revistas),” IEEE Transl. J. Magn.Jpn., vol. 2, Aug. 1987, pp. 740–741 [Dig. 9th Annu. Conf. Magnetics Japan, 1982, p. 301].
17. M. Young, The Techincal Writers Handbook. Mill Valley, CA: University Science, 1989.
18. J. U. Duncombe, “Infrared navigation—Part I: An assessment of feasibility (Estilo para revistas especializadas),” IEEE Trans. Electron Devices, vol. ED-11, pp. 34–39, Jan. 1959.
19. S. Chen, B. Mulgrew, and P. M. Grant, “A clustering technique for digital communications channel equalization using radial basis function networks,” IEEE Trans. Neural Networks, vol. 4, pp. 570–578, July 1993.
20. R. W. Lucky, “Automatic equalization for digital communication,” Bell Syst. Tech. J., vol. 44, no. 4, pp. 547–588, Apr. 1965.
21. S. P. Bingulac, “On the compatibility of adaptive controllers (Estilo para memorias de congresos publicadas),” in Proc. 4th Annu. Allerton Conf. Circuits and Systems Theory, New York, 1994, pp. 8–16.
22. G. R. Faulhaber, “Design of service systems with priority reservation,” in Conf. Rec. 1995 IEEE Int. Conf. Communications, pp. 3–8.
23. W. D. Doyle, “Magnetization reversal in films with biaxial anisotropy,” in 1987 Proc. INTERMAG Conf., pp. 2.2-1–2.2-6.
24. G. W. Juette and L. E. Zeffanella, “Radio noise currents n short sections on bundle conductors (Estilo para trabajos presentados en congresos),” presented at the IEEE Summer power Meeting, Dallas, TX, June 22–27, 1990, Paper 90 SM 690-0 PWRS.
25. J. G. Kreifeldt, “An analysis of surface-detected EMG as an amplitude-modulated noise,” presented at the 1989 Int. Conf. Medicine and Biological Engineering, Chicago, IL.
26. J. Williams, “Narrow-band analyzer (Estilo para Tesis),” Ph.D. dissertation, Dept. Elect. Eng., Harvard Univ., Cambridge, MA, 1993.
27. N. Kawasaki, “Parametric study of thermal and chemical nonequilibrium nozzle flow,” M.S. thesis, Dept. Electron. Eng., Osaka Univ., Osaka, Japan, 1993.
28. J. P. Wilkinson, “Nonlinear resonant circuit devices (Estilo para patentes),” U.S. Patent 3 624 12, July 16, 1990.
29. IEEE Criteria for Class IE Electric Systems (Estilo para estándares), IEEE Standard 308, 1969.
30. Letter Symbols for Quantities, ANSI Standard Y10.5-1968.
31. R. E. Haskell and C. T. Case, “Transient signal propagation in lossless isotropic plasmas (Estilo para reportes),” USAF Cambridge Res. Lab., Cambridge, MA Rep. ARCRL-66-234 (II), 1994, vol. 2.
32. E. E. Reber, R. L. Michell, and C. J. Carter, “Oxygen absorption in the Earth’s atmosphere,” Aerospace Corp., Los Angeles, CA, Tech. Rep. TR-0200 (420-46)-3, Nov. 1988.
33. (Estilo para manuales) Transmission Systems for Communications, 3rd ed., Western Electric Co., Winston-Salem, NC, 1985, pp. 44–60.
34. Motorola Semiconductor Data Manual, Motorola Semiconductor Products Inc., Phoenix, AZ, 1989.
35. (Estilo para libros y monografías en Internet) J. K. Author. (year, month, day). Title (edition) [Type of medium]. Volume(issue). Available: http://www.(URL)
36. J. Jones. (1991, May 10). Networks (2nd ed.) [Online]. Available: http://www.atm.com
37. (Estilo para revistas electrónicas) K. Author. (year, month). Title. Journal [Type of medium]. Volume(issue), paging if given. Available: http://www.(URL)
38. R. J. Vidmar. (1992, August). On the use of atmospheric plasmas as electromagnetic reflectors. IEEE Trans. Plasma Sci. [Online]. 21(3). pp. 876—880. Available: http://www.halcyon.com/pub/journals/21ps03-vidmar

**8. Alumnos y Directores**

CARÁCTER: Confidencial

FUNDAMENTO LEGAL: Art. 3, fracc. II, Art. 18, fracc. II y Art. 21, lineamiento 32, fracc. XVII de la L.F.T.A.I.P.G.

PARTES CONFIDENCIALES: No. de boleta y Teléfono.

|  |
| --- |
| *Nombre del alumno*.- Alumno de la carrera de Nombre de la Carrera en Nombre de la Institución, Especialidad Nombre de la Especialidad, Boleta: , Tel. , email |

Firma:\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

TURNO PARA LA PRESENTACIÓN DEL TRABAJO TERMINAL:

|  |
| --- |
| *Marco Antonio Fernández Ramírez*.- Alumno de la carrera de Ing. en Sistemas Computacionales en ESCOM, Especialidad Sistemas, Boleta: , Tel. , email m\_a\_f\_r@hotmail.com. |

Firma:\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

|  |
| --- |
| *Nombre del Director*.- Breve currículum enfatizando su capacidad para dirigir el TT, Áreas de Interés, Datos de contacto (Teléfono, email). |

Firma:\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

|  |
| --- |
| *Acosta Bermejo Raúl*.- Dr. en Informática, Tiempo Real, Robótica y Automatismo del École de Mines de Pais en 2003, M. en C. en Computación del CINVESTAV en 1997, Ing. en Electrónica de la UAM en 1993, Profesor de ESCOM/IPN (Dpto de Posgrado) desde 1993, Áreas de Interés: MRS, Redes. Ext. 52028, email racostab@ipn.mx. |

Firma:\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Nombre del alumno(a): \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ TT No.:

Título del TT:

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Actividad | AGO | SEP | OCT | NOV | DIC | ENE | FEB | MAR | ABR | MAY |
| Análisis y diseño del sistema |  |  |  |  |  |  |  |  |  |  |
| Evaluación de TT I. |  |  |  |  |  |  |  |  |  |  |
| Generación del código. |  |  |  |  |  |  |  |  |  |  |
| Pruebas. |  |  |  |  |  |  |  |  |  |  |
| Reingeniería. |  |  |  |  |  |  |  |  |  |  |
| Generación del Manual de Usuario y la Página web. |  |  |  |  |  |  |  |  |  |  |
| Generación el Reporte Técnico. |  |  |  |  |  |  |  |  |  |  |
| Presentar los resultados en congresos. |  |  |  |  |  |  |  |  |  |  |
| Evaluación de TT II. |  |  |  |  |  |  |  |  |  |  |