



Casa abierta al tiempo

**UNIVERSIDAD AUTÓNOMA METROPOLITANA**  
Unidad Cuajimalpa

*Comunidad académica comprometida  
con el desarrollo humano de la sociedad.*

DCCD.DTI.021.19  
Mayo 20, 2019

**Dr. Octavio Mercado González**

Presidente del Consejo Divisional de la  
División de Ciencias de la Comunicación y Diseño  
Unidad Cuajimalpa  
Presente

**Asunto:** Solicitud de Periodo Sabático del  
Dr. Héctor Jiménez Salazar

*Estimado Dr. Mercado:*

Con relación al asunto arriba referido, por este conducto me permito solicitarle se someta a consideración del Consejo Divisional la petición que el Dr. Héctor Jiménez Salazar me ha enviado para disfrutar de un Periodo Sabático de seis meses, que iniciaría en septiembre de 2019.

Para tal efecto, anexo copia de la documentación que me hizo llegar la Coordinación de Recursos Humanos de esta Unidad, en la cual se hace constar que el Dr. Héctor Jiménez Salazar ha cumplido los requisitos de tiempo de antigüedad para disfrutar de tal periodo sabático. Anexo también el Plan de Trabajo que presenta el Dr. Héctor Jiménez Salazar.

Se envían los documentos indicados y anexados en formato digital, vía correo electrónico.

Sin otro particular, envío a usted un cordial saludo.

**Atentamente,**

Casa abierta al tiempo

**Dr. Carlos Joel Rivero Moreno**

Jefe del Departamento de Tecnologías de la Información

Anexo: Lo mencionado.

c.c.p.: Dra. Gloria Angélica Martínez de la Peña – Secretaria del Consejo Divisional  
Lic. Inés Andrea Zepeda Martínez – Oficina Técnica de Consejo Divisional.

CJRM



**Unidad Cuajimalpa**

DCCD | Jefatura del Departamento de Tecnologías de la Información  
Torre III, 5to. piso. Avenida Vasco de Quiroga 4871,  
Colonia Santa Fe Cuajimalpa. Delegación Cuajimalpa de Morelos,  
Tel. +52 (55) 5814-6557, C.P. 05348, México, D.F.  
<http://dccd.cua.uam.mx>



Casa abierta al tiempo

**UNIVERSIDAD AUTÓNOMA METROPOLITANA**  
Unidad Cuajimalpa

*Comunidad académica comprometida  
con el desarrollo humano de la sociedad.*

**RHC.081.2019**

**PERIODO SABÁTICO**

**14 de mayo 2019**

**DR. HÉCTOR JIMÉNEZ SALAZAR (31898)**  
**DEPARTAMENTO DE TECNOLOGÍAS DE LA INFORMACIÓN**  
**DIVISIÓN DE CIENCIAS DE LA COMUNICACIÓN Y DISEÑO**  
**UNIDAD CUAJIMALPA**  
Presente.

**Estimado Dr. Jiménez**

Conforme a su petición y de acuerdo a nuestros registros y a su trayectoria laboral dentro de nuestra Institución, usted inicia la acumulación de tiempo para el disfrute de periodo sabático, a partir del 1 de marzo de 2007 y durante su estancia laboral en esta Universidad, no ha disfrutado de ningún periodo sabático.

Para esta fecha usted acumula para periodo sabático, doce años, dos meses, trece días, de labores ininterrumpidas en su plaza académica al servicio de la Universidad, por lo que puede solicitar y disfrutar de un periodo sabático por un tiempo máximo hasta de veinticuatro meses (dos años).

Sin otro particular, estoy a sus apreciables órdenes para cualquier aclaración al respecto.



**ATENTAMENTE**

**"CASA ABIERTA AL TIEMPO"**

Casa abierta al tiempo

**UNIVERSIDAD AUTÓNOMA METROPOLITANA**  
**UNIDAD CUAJIMALPA**

**RECURSOS HUMANOS**

**LIC. LUIS BECERRA CASTAÑEDA**  
**COORDINADOR DE RECURSOS HUMANOS**

**C.c.p. Mtro. Octavio Mercado González, Presidente del Consejo Divisional DCCD.**  
**Dr. Carlos Joel Rivero Moreno, Jefe del Depto. de Tecnologías de la Información, DCCD.**  
**Expediente**

**Unidad Cuajimalpa**

Secretaría de Unidad

Torre III, 8to. piso. Avenida Vasco de Quiroga 4871, Colonia Santa Fe Cuajimalpa

Delegación Cuajimalpa de Morelos, CDMX., C.P. 05348,

Tel. 5814-6505 a 07; correo electrónico: cgarcia@correo.cua.uam.mx

www.cua.uam.mx

# Plan de actividades para un posible ejercicio de periodo sabático de seis meses

Héctor Jiménez Salazar  
DTI, DCCD, UAM-C

Mayo 2019

## 1 Introducción

Somos testigos de un caudal de software cuyo incremento reciente ha sido de gran impacto en los más variados contextos. Particularmente, software que atiende el problema de la ‘predicción’, es decir, aquél que con base en las regularidades de cierto fenómeno, propone el comportamiento de éste en nuevas circunstancias [3]. Los métodos subyacentes en estas aplicaciones se conciben ‘entrenando’, con un conjunto de ejemplos resueltos (descripciones de problemas y su solución) un modelo que concentra la forma de tomar decisiones frente a nuevos ejemplares de los cuales se desconoce la solución. En suma, dichos sistemas realizan la clasificación de nuevos ejemplares con base en lo antes ‘aprendido’. Si bien este tipo de herramientas se ha venido desarrollando desde los inicios de la estadística [1], su expansión a nuevos campos de aplicación se ha intensificado a través de algunas áreas de la computación como el aprendizaje computacional. Aún con todo este desarrollo, queda por atender el problema de la selección de las características de las instancias para poder ser procesadas con los diversos métodos de clasificación; la principal razón de esta selección es representar de manera adecuada los ejemplares para entrenar el modelo y evitar, con la eliminación de características espurias, que se introduzca ruido en el proceso. La selección de características, o selección de atributos (*Feature Selection*) es aún un problema abierto [2] que requiere explorar el alcance de los métodos, su desempeño y el contexto de aplicación para poder ofrecer métodos de clasificación más robustos.

En este plan de actividades se pretende ubicar el desempeño de un método de selección de atributos dentro de la constelación de técnicas existentes cruzadas con los tipos de datos, colecciones de diferente naturaleza, y su prueba en los clasificadores más utilizados. Las secciones restantes de este documento exponen el objetivo, las actividades y las metas de este plan.

## 2 Selección de atributos

En la literatura hay algunas referencias al empleo de métodos que, en un espacio multidimensional [5] (dimensión dada por el número de atributos de las instancias) consideran hiperplanos de separación en el problema de clasificación binaria. La construcción de tales hiperplanos trata de descubrir las fronteras entre los ejemplares representados en el espacio si existe separación entre las clases. El problema de clases binarias, clases separables y linealidad es el más sencillo y es donde se empezará a explorar. La simple idea de calcular la distancia entre los centroides de las clases, para averiguar si estos se acercan o distancian al eliminar un atributo de las instancias, se espera ir adaptando a nuevas condiciones; por ejemplo determinar todas las distancias entre cada par de centroides en un problema multiclase.

Hay diversos métodos de selección de atributos, como ganancia de información (GI),  $\chi^2$ , o el empleo de la varianza de los atributos para tomar la decisión sobre cuáles de ellos son más influyentes en la clasificación. Por un lado se trata de realizar experimentos con la separación de centroides y comparar los resultados con los métodos de selección más comunes. Ya que se espera evaluar el desempeño de cada método realizando la clasificación y así conocer la efectividad de la selección, deberán elegirse algunos métodos de clasificación. También, en este caso, se considera trabajar con los métodos más utilizados, por ejemplo, árboles de decisión, K-nn, y Bayes. Por último, puesto que los dominios para llevar a cabo la clasificación son muy variados, deberán tomarse colecciones de datos que puedan influir en la clasificación con el fin de limitar o ampliar el alcance de los métodos de selección. En este último caso se trabajará con corpora de textos, principalmente textos cortos, tratando de comparar los resultados obtenidos con otros métodos reportados en la literatura.

### 2.1 Objetivo

Conocer el desempeño del método de selección de atributos basado en la separación de centroides en diversas colecciones de datos, asimismo, confrontar su desempeño con otros métodos de selección de atributos a través de la clasificación de instancias de cada colección con los clasificadores clásicos (Bayes, árboles de decisión, y K-nn).

### 2.2 Cronograma

Como se ha mencionado, las actividades son eminentemente experimentales tomando como base algunos resultados obtenidos previamente [4] que permiten guiar la exploración del método en nuevas circunstancias. Por lo anterior se propone el siguiente cronograma.

Periodo	Actividad
Mes 1	Revisión de trabajos relacionados y obtención de corpora
Mes 2	Experimentos para conocer desempeño en los corpora
Mes 3	Replicación de algunos métodos, comparación y análisis de resultados
Mes 4	Escritura de reporte para publicación
Mes 5	Experimentos con variantes del método y análisis de resultados
Mes 6	Reporte final

Se espera realizar las actividades mencionadas en colaboración con algunos miembros de la AMPLN (Asociación Mexicana del Procesamiento del Lenguaje Natural) quienes pertenecen a otras instituciones; INAOE, CIC-IPN, y CIMAT.

### 2.3 Metas

De acuerdo con el objetivo se proponen las siguientes metas:

1. Publicación de un artículo sobre la comparación del método basado en separación de centroides con otros métodos de selección de atributos.
2. Un reporte sobre la continuidad de la investigación realizada; por ejemplo, cuáles son los problemas principales y posibles maneras de atender para mejorar desempeño.

## Referencias

- [1] Fisher, R. A. (1936). The use of multiple measurements in taxonomic problems. *Annals of eugenics*, 7(2), 179-188.
- [2] Guyon, I., & Elisseeff, A. (2003). An introduction to variable and feature selection. *Journal of machine learning research*, 3(Mar), 1157-1182.
- [3] David J. Hand, Padhraic Smyth, Heikki Mannila (2001). *Principles of data mining*, MIT Press Cambridge, MA, USA.
- [4] Héctor Jiménez-Salazar, Alfredo Mateos-Papis, Christian Sánchez-Sánchez (2019). Selección de atributos mediante separación de centroides. Aceptado para su publicación en *Research in Computing Science*.
- [5] Salton, G., Wong, A., & Yang, C. S. (1975). A vector space model for automatic indexing. *Communications of the ACM*, 18(11), 613-620.