

Gestión de Conocimiento a través del Análisis de las Redes Sociales y la Información Contenida en los Correos Electrónicos Personales Estado del Arte

Juan David Cruz Gómez
299618

Maestría en Ingeniería – Ingeniería de Sistemas
Universidad Nacional de Colombia

Seminario de Investigación I

Agenda

- 1 Introducción
 - Definición del problema
- 2 Gestión de Conocimiento
- 3 Análisis de Redes Sociales
- 4 Extracción de Conocimiento
 - Aproximación Determinística
 - Aproximaciones Probabilísticas
 - Aproximación Bio-inspirada
- 5 Preguntas

Agenda

- 1 **Introducción**
 - Definición del problema
- 2 Gestión de Conocimiento
- 3 Análisis de Redes Sociales
- 4 Extracción de Conocimiento
 - Aproximación Determinística
 - Aproximaciones Probabilísticas
 - Aproximación Bio-inspirada
- 5 Preguntas

Algunos Hechos

- Masificación en el desarrollo y uso de las tecnologías de la información y comunicaciones.
- Crecimiento en la utilización del correo electrónico como canal de comunicación rápido para el intercambio de información.
- El volumen diario de correos entrantes tiende a ser inmanejable.
- Dentro de los mensajes de correo existe implícito un conocimiento del remitente y tiene algún tema asociado.
- Los mensajes de correo electrónico de forma natural, pueden generar una red de personas.

Algunos Hechos

- Masificación en el desarrollo y uso de las tecnologías de la información y comunicaciones.
- Crecimiento en la utilización del correo electrónico como canal de comunicación rápido para el intercambio de información.
- El volumen diario de correos entrantes tiende a ser inmanejable.
- Dentro de los mensajes de correo existe implícito un conocimiento del remitente y tiene algún tema asociado.
- Los mensajes de correo electrónico de forma natural, pueden generar una red de personas.

Algunos Hechos

- Masificación en el desarrollo y uso de las tecnologías de la información y comunicaciones.
- Crecimiento en la utilización del correo electrónico como canal de comunicación rápido para el intercambio de información.
- El volumen diario de correos entrantes tiende a ser inmanejable.
- Dentro de los mensajes de correo existe implícito un conocimiento del remitente y tiene algún tema asociado.
- Los mensajes de correo electrónico de forma natural, pueden generar una red de personas.

Algunos Hechos

- Masificación en el desarrollo y uso de las tecnologías de la información y comunicaciones.
- Crecimiento en la utilización del correo electrónico como canal de comunicación rápido para el intercambio de información.
- El volumen diario de correos entrantes tiende a ser inmanejable.
- Dentro de los mensajes de correo existe implícito un conocimiento del remitente y tiene algún tema asociado.
- Los mensajes de correo electrónico de forma natural, pueden generar una red de personas.

Algunos Hechos

- Masificación en el desarrollo y uso de las tecnologías de la información y comunicaciones.
- Crecimiento en la utilización del correo electrónico como canal de comunicación rápido para el intercambio de información.
- El volumen diario de correos entrantes tiende a ser inmanejable.
- Dentro de los mensajes de correo existe implícito un conocimiento del remitente y tiene algún tema asociado.
- Los mensajes de correo electrónico de forma natural, pueden generar una red de personas.

El Problema...

- El conocimiento contenido en los mensajes de correo no es almacenado correctamente y generalmente se pierde.
- Como relacionar los documentos y los correos personales de un individuo
- Como relacionar el conocimiento explícito de documentos con el conocimiento tácito de las redes sociales.

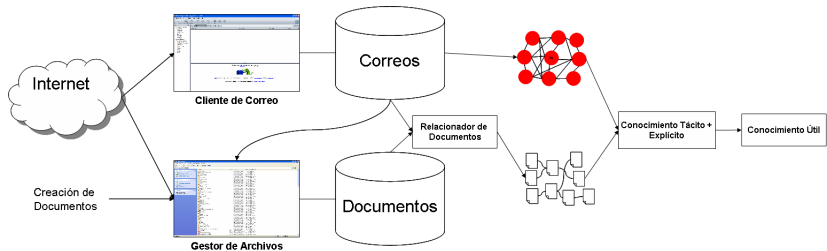
El Problema...

- El conocimiento contenido en los mensajes de correo no es almacenado correctamente y generalmente se pierde.
- Como relacionar los documentos y los correos personales de un individuo
- Como relacionar el conocimiento explícito de documentos con el conocimiento tácito de las redes sociales.

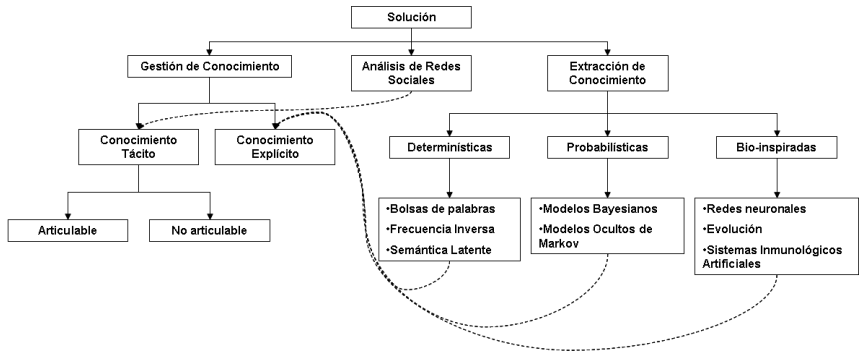
El Problema...

- El conocimiento contenido en los mensajes de correo no es almacenado correctamente y generalmente se pierde.
- Como relacionar los documentos y los correos personales de un individuo
- Como relacionar el conocimiento explícito de documentos con el conocimiento tácito de las redes sociales.

El Problema...



Taxonomía



Gestión de Conocimiento

- Es la gestión del capital intelectual en una organización, con el fin de agregar valor a los productos y servicios ofrecidos.
- El conocimiento puede ser dividido en dos tipos: tácito y explícito.

Gestión de Conocimiento

- Es la gestión del capital intelectual en una organización, con el fin de agregar valor a los productos y servicios ofrecidos.
- El conocimiento puede ser dividido en dos tipos: tácito y explícito.

Conocimiento Tácito

- El conocimiento tácito o implícito es aquel que poseen las personas y que no puede ser fácilmente extraído.
- Gestionar este conocimiento es bastante complicado.
- Se divide en dos: articulable y no articulable.

Conocimiento Tácito

- El conocimiento tácito o implícito es aquel que poseen las personas y que no puede ser fácilmente extraído.
- Gestionar este conocimiento es bastante complicado.
- Se divide en dos: articulable y no articulable.

Conocimiento Tácito

- El conocimiento tácito o implícito es aquel que poseen las personas y que no puede ser fácilmente extraído.
- Gestionar este conocimiento es bastante complicado.
- Se divide en dos: articulable y no articulable.

Conocimiento Explícito

- El conocimiento explícito es todo aquel conocimiento que está disponible en documentos, bases de datos o cual otro tipo de repositorio físico, que además, describe algún procedimiento o técnica. Esto quiere decir que puede ser aprendido a través de libros o de artículos y es fácilmente extraíble y administrable.

¿Qué son las Redes Sociales?

- Las redes sociales son representaciones de las interacciones entre personas o entre organizaciones.
- La representación es forma de un grafo dirigido, en donde los nodos corresponden a individuos (personas, organizaciones o cualquier asociación) y los arcos son los vínculos entre ellos.
- La teoría que rodea a las redes sociales es utilizada para analizar los vínculos entre individuos.

¿Qué son las Redes Sociales?

- Las redes sociales son representaciones de las interacciones entre personas o entre organizaciones.
- La representación es forma de un grafo dirigido, en donde los nodos corresponden a individuos (personas, organizaciones o cualquier asociación) y los arcos son los vínculos entre ellos.
- La teoría que rodea a las redes sociales es utilizada para analizar los vínculos entre individuos.

¿Qué son las Redes Sociales?

- Las redes sociales son representaciones de las interacciones entre personas o entre organizaciones.
- La representación es forma de un grafo dirigido, en donde los nodos corresponden a individuos (personas, organizaciones o cualquier asociación) y los arcos son los vínculos entre ellos.
- La teoría que rodea a las redes sociales es utilizada para analizar los vínculos entre individuos.

Formas de Realizar Mediciones en Redes Sociales

- **Grado:** Un individuo con alto grado, tiene más posibilidades de satisfacer sus necesidades y se reduce la dependencia con otros individuos.
- **Intermediación:** Un individuo con alto grado de intermediación, se convierte en indispensable y al mismo tiempo se convierte en un punto vulnerable de la red, ya que puede generar huecos estructurales que desconectan otros individuos o redes completas.
- **Cercanía:** Los individuos con un alto grado de cercanía, tienen a su disposición los caminos más cortos hacia otros individuos, por lo que se encuentran en una excelente posición para monitorear el flujo de información que fluye a través de esta.

Formas de Realizar Mediciones en Redes Sociales

- **Grado:** Un individuo con alto grado, tiene más posibilidades de satisfacer sus necesidades y se reduce la dependencia con otros individuos.
- **Intermediación:** Un individuo con alto grado de intermediación, se convierte en indispensable y al mismo tiempo se convierte en un punto vulnerable de la red, ya que puede generar huecos estructurales que desconectan otros individuos o redes completas.
- **Cercanía:** Los individuos con un alto grado de cercanía, tienen a su disposición los caminos más cortos hacia otros individuos, por lo que se encuentran en una excelente posición para monitorear el flujo de información que fluye a través de esta.

Formas de Realizar Mediciones en Redes Sociales

- Grado: Un individuo con alto grado, tiene más posibilidades de satisfacer sus necesidades y se reduce la dependencia con otros individuos.
- Intermediación: Un individuo con alto grado de intermediación, se convierte en indispensable y al mismo tiempo se convierte en un punto vulnerable de la red, ya que puede generar huecos estructurales que desconectan otros individuos o redes completas.
- Cercanía: Los individuos con un alto grado de cercanía, tienen a su disposición los caminos más cortos hacia otros individuos, por lo que se encuentran en una excelente posición para monitorear el flujo de información que fluye a través de esta.

Agenda

- 1 Introducción
 - Definición del problema
- 2 Gestión de Conocimiento
- 3 Análisis de Redes Sociales
- 4 Extracción de Conocimiento**
 - **Aproximación Determinística**
 - Aproximaciones Probabilísticas
 - Aproximación Bio-inspirada
- 5 Preguntas

Vectorización de Textos

- Se busca la frecuencia de cada una de las palabras en un texto y cada una de dichas frecuencias corresponderá a un valor, normalizado, en el eje determinado por dicha palabra.
- El manejo de documentos utilizando esta técnica puede ser bastante complejo.
- Todas las palabras tienen la misma importancia o peso.

Vectorización de Textos

- Se busca la frecuencia de cada una de las palabras en un texto y cada una de dichas frecuencias corresponderá a un valor, normalizado, en el eje determinado por dicha palabra.
- El manejo de documentos utilizando esta técnica puede ser bastante complejo.
- Todas las palabras tienen la misma importancia o peso.

Vectorización de Textos

- Se busca la frecuencia de cada una de las palabras en un texto y cada una de dichas frecuencias corresponderá a un valor, normalizado, en el eje determinado por dicha palabra.
- El manejo de documentos utilizando esta técnica puede ser bastante complejo.
- Todas las palabras tienen la misma importancia o peso.

Agenda

- 1 Introducción
 - Definición del problema
- 2 Gestión de Conocimiento
- 3 Análisis de Redes Sociales
- 4 Extracción de Conocimiento**
 - Aproximación Determinística
 - Aproximaciones Probabilísticas**
 - Aproximación Bio-inspirada
- 5 Preguntas

Técnicas Bayesianas

- Estas técnicas se basan en la siguiente supocisión: se tienen m clases de tópicos, $c_1, c_2, \dots, c_m$ con algunos documentos semilla D_c asociados a cada una donde, la probabilidad de que un documento pertenezca a una clase está dado por : $\frac{|D_c|}{\sum_c |D_c|}$.
- Se define la probabilidad de que un documento d pertenezca a una clase c , $\Pr(d | c)$ como:

$$\arg \max \left(\Pr(d | c) = \prod_{t \in d} \phi_{c,t} \prod_{t \in T, t \notin d} (1 - \phi_{c,t}) \right) \text{ donde } T$$

es el universo o vocabulario de términos y $\phi_{c,t}$, el cual indica la probabilidad de que un documento en la clase c contenga el término t al menos una vez.

Modelos Ocultos de Markov

- En un modelo de Markov solo interesa el valor del estado inmediatamente anterior, el modelo oculto de Markov se entrena para encontrar la matriz $\mathbf{P}$ y el vector $\mathbf{B}$, de modo que se pueda encontrar el camino más probable entre los estados dado un símbolo de entrada.
- Cada uno de los estados o clases está en capacidad de emitir uno de los términos del vocabulario con cierta probabilidad.

Agenda

- 1 Introducción
 - Definición del problema
- 2 Gestión de Conocimiento
- 3 Análisis de Redes Sociales
- 4 Extracción de Conocimiento
 - Aproximación Determinística
 - Aproximaciones Probabilísticas
 - Aproximación Bio-inspirada
- 5 Preguntas

Redes Neuronales

- Estas técnicas simulan el comportamiento del sistema nervioso central.
- Existen dos tipos importantes: las redes neuronales artificiales y los mapas de Kohonen o mapas auto organizativos.
- Las redes neuronales artificiales sirven principalmente para realizar clasificación.
- Los mapas auto-organizativos son utilizados para realizar tareas de agrupamiento.

Redes Neuronales

- Estas técnicas simulan el comportamiento del sistema nervioso central.
- Existen dos tipos importantes: las redes neuronales artificiales y los mapas de Kohonen o mapas auto organizativos.
- Las redes neuronales artificiales sirven principalmente para realizar clasificación.
- Los mapas auto-organizativos son utilizados para realizar tareas de agrupamiento.

Redes Neuronales

- Estas técnicas simulan el comportamiento del sistema nervioso central.
- Existen dos tipos importantes: las redes neuronales artificiales y los mapas de Kohonen o mapas auto organizativos.
- Las redes neuronales artificiales sirven principalmente para realizar clasificación.
- Los mapas auto-organizativos son utilizados para realizar tareas de agrupamiento.

Redes Neuronales

- Estas técnicas simulan el comportamiento del sistema nervioso central.
- Existen dos tipos importantes: las redes neuronales artificiales y los mapas de Kohonen o mapas auto organizativos.
- Las redes neuronales artificiales sirven principalmente para realizar clasificación.
- Los mapas auto-organizativos son utilizados para realizar tareas de agrupamiento.

Sistemas Inmunológicos Artificiales

- Basan su funcionamiento en el modelo de defensa corporal.
- La primera aproximación es basada en el principio propio – extraño.
- Selección clonal, redes inmunes y memoria inmune.
- Se generan detectores para realizar tareas de clasificación y agrupamiento.

Sistemas Inmunológicos Artificiales

- Basan su funcionamiento en el modelo de defensa corporal.
- La primera aproximación es basada en el principio propio – extraño.
- Selección clonal, redes inmunes y memoria inmune.
- Se generan detectores para realizar tareas de clasificación y agrupamiento.

Sistemas Inmunológicos Artificiales

- Basan su funcionamiento en el modelo de defensa corporal.
- La primera aproximación es basada en el principio propio – extraño.
- Selección clonal, redes inmunes y memoria inmune.
- Se generan detectores para realizar tareas de clasificación y agrupamiento.

Sistemas Inmunológicos Artificiales

- Basan su funcionamiento en el modelo de defensa corporal.
- La primera aproximación es basada en el principio propio – extraño.
- Selección clonal, redes inmunes y memoria inmune.
- Se generan detectores para realizar tareas de clasificación y agrupamiento.

Computación Evolutiva

- Se basa en la teoría de la evolución de Darwin.
- Busca evolucionar poblaciones para encontrar el individuo mejor adaptado a un problema particular.
- Buscando optimizar alguna función, se pueden generar agrupaciones o clasificaciones de documentos.

Computación Evolutiva

- Se basa en la teoría de la evolución de Darwin.
- Busca evolucionar poblaciones para encontrar el individuo mejor adaptado a un problema particular.
- Buscando optimizar alguna función, se pueden generar agrupaciones o clasificaciones de documentos.

Computación Evolutiva

- Se basa en la teoría de la evolución de Darwin.
- Busca evolucionar poblaciones para encontrar el individuo mejor adaptado a un problema particular.
- Buscando optimizar alguna función, se pueden generar agrupaciones o clasificaciones de documentos.

Preguntas de Investigación

- 1 ¿Cómo relacionar el conocimiento explícito con el conocimiento tácito incluido en una red social?
- 2 ¿Que medidas de las redes sociales debo utilizar?
- 3 ¿Cuál o cuales métodos de extracción de conocimiento de documentos es el mejor para este tipo de problemas?
- 4 ¿Es factible realizar recomendaciones de colaboración en una red social personal?

Preguntas de Investigación

- 1 ¿Cómo relacionar el conocimiento explícito con el conocimiento tácito incluido en una red social?
- 2 ¿Que medidas de las redes sociales debo utilizar?
- 3 ¿Cuál o cuales métodos de extracción de conocimiento de documentos es el mejor para este tipo de problemas?
- 4 ¿Es factible realizar recomendaciones de colaboración en una red social personal?

Preguntas de Investigación

- 1 ¿Cómo relacionar el conocimiento explícito con el conocimiento tácito incluido en una red social?
- 2 ¿Que medidas de las redes sociales debo utilizar?
- 3 ¿Cuál o cuales métodos de extracción de conocimiento de documentos es el mejor para este tipo de problemas?
- 4 ¿Es factible realizar recomendaciones de colaboración en una red social personal?

Preguntas de Investigación

- 1 ¿Cómo relacionar el conocimiento explícito con el conocimiento tácito incluido en una red social?
- 2 ¿Que medidas de las redes sociales debo utilizar?
- 3 ¿Cuál o cuales métodos de extracción de conocimiento de documentos es el mejor para este tipo de problemas?
- 4 ¿Es factible realizar recomendaciones de colaboración en una red social personal?