



UNIVERSIDAD DE CONCEPCIÓN
FACULTAD DE INGENIERÍA
PROGRAMA DE MAGÍSTER EN CIENCIAS DE LA COMPUTACIÓN

Evaluación de Reglas de Asociación en Text Mining Utilizando Métricas Semánticas y Estructurales



Profesor Guía: John Atkinson Abutridy
Dpto. de Informática y Ciencias de la Computación
Facultad de Ingeniería
Universidad de Concepción

Tesis para ser presentada a la Dirección de Postgrado de la Universidad
de Concepción

CLAUDIO ALBERTO PÉREZ CÁRCAMO
CONCEPCIÓN-CHILE
2010

Resumen

Tareas tradicionales en minería de textos descubren patrones desconocidos considerados *interesantes* desde un punto de vista estadístico usando métricas estándar. Esta evaluación estadística no toma en cuenta modelos mentales, usuarios o conocimiento del dominio. Mientras existen algunos enfoques para evaluar patrones (reglas de asociación) en text mining éstos fallan en determinar con exactitud cuales de los patrones descubiertos resultan ser *interesantes/novedosos*. En esta tesis, se propone y discute un nuevo enfoque que combina semántica basada en corpus (*LSA*) y modelo de conocimiento basado en Lattices. Los Experimentos realizados utilizando nuestro enfoque en comparación con otros modelos de evaluación del estado del arte muestran lo promisorio de nuestro método para evaluar los patrones descubiertos como *interesantes* o *sin interés* hacia expertos humanos.

