

KLASIFIKASI TUGAS AKHIR BERDASARKAN JURNAL DAN ONTOLOGI TUGAS AKHIR STUDI KASUS : TUGAS AKHIR MAHASISWA JURUSAN TEKNIK INFORMATIKA STT TELKOM NAÏVE BAYES CLASSIFIER IMPLEMENTATION FOR FINAL PROJECT CLASSIFICATION BASED ON JOURNAL AND FINAL PROJECT ONTOLOG

Muhamad Rifki Setyadji^{1, -2}

¹Teknik Informatika, Fakultas Teknik Informatika, Universitas Telkom

Abstrak

Akibat dari belum adanya sistem yang menyimpan dan melakukan proses manajemen terhadap arsip tugas akhir (TA) di jurusan teknik informatika STT Telkom selama ini, muncul beberapa permasalahan antara lain berhubungan dengan proses query terhadap arsip yang ada, standar klasifikasi yang tidak kunjung mapan, dan sulitnya mencari keterkaitan antara TA yang satu dengan lainnya. Namun, dengan melihat fakta bahwa judul yang ada mencapai ratusan, sebaiknya proses klasifikasi arsip yang ada tidak dilakukan secara manual.

Pada tugas akhir ini, dihasilkan sebuah perangkat lunak manajemen TA menggunakan ontologi, sebagai bentuk representasi formal yang dipilih karena dipandang mampu menjawab permasalahan manajemen tugas akhir mahasiswa jurusan Teknik Informatika STT Telkom. Proses klasifikasi akan berjalan secara otomatis dengan menggunakan metoda Naïve Bayes Classifier. Uji coba yang dilakukan adalah untuk mengukur akurasi dan tingkat kesalahan yang muncul, bukan untuk dibandingkan dengan metode yang berbeda. Kakas yang dipakai adalah Natural Language Toolkit (NLTK). Dengan memanfaatkan modul NaiveBayesClassifier yang ada didalamnya, dihasilkan sebuah klasifikator dokumen tugas akhir.

Kata Kunci : naïve bayes classifier, ontologi, manajemen tugas akhir

Abstract

Since we haven't yet system that manage the final project so far, then I have an idea for that. With this final project, this management mechanism should work effectively and substitute the existing mechanism.

Ontology, as formal representation, chosen as a system that we hope can answer the final project management related problem. The ontology structure must be connected to the rules of informatics department on STT Telkom appropriately.

As fact, we have about hundreds title of final project on the informatics department. We must develop a new mechanism to make it efficient while input the data to the ontology database. We must avoid that process manually. But, maybe the result will not too accurate when we compare with manual process.

Naïve Bayes Classifier will be function as tool for extracting information or keywords on the final project document (journal) archives. This scheme can manipulate it to filling the ontologi structure automatically.

This final project won't show that Naïve Bayes Classifier is the best mechanism for document extraction. It just want to prove that we can use the Naïve Bayes Classifier for extracting some keywords from text document or not. If available, then the application will generate an output that figures accuraction and error level.

Keywords : naïve bayes classifier, ontology, final project management

BAB I

PENDAHULUAN

1.1 Latar Belakang

Jurusan Teknik Informatika STT Telkom yang telah ada sejak 15 tahun lalu memiliki koleksi ratusan dokumen tugas akhir. Koleksi tersebut bahkan tidak akan berhenti sampai saat ini saja. Dalam hal ini, manajemen pengetahuan sangat diperlukan untuk menjalankan fungsinya; distribusi, organisasi, prosedur, dan pengaksesan sekaligus memantau dan mengendalikan realisasi dari pola topik-topik yang diteliti.

Untuk membentuk sebuah sistem manajemen pengetahuan, perlu direpresentasikan ke dalam sebuah bentuk pengelolaan yang memang berbasis pengetahuan, salah satunya adalah ontologi. Ontologi merupakan satu cara untuk mengklasifikasikan objek-objek yang ada di dalam kumpulannya. Dalam hal ini, kumpulan tugas akhir akan diklasifikasi berdasarkan struktur ontologi.

Untuk proses pengisian ontologi tugas akhir, dibutuhkan proses klasifikasi yang berjalan secara otomatis. Hal ini perlu dipikirkan karena seiring dengan berjalannya waktu, dokumen yang diproses akan semakin banyak dan sangat tidak efisien jika pengklasifikasian harus dilakukan secara manual. Untuk itu penulis mencoba memakai salah satu metode klasifikasi, yaitu dengan menggunakan naïve bayes classifier. Sekali lagi metoda ini hanyalah salah satu dari sekian banyak metode klasifikasi yang lain. Alasan penggunaan metode ini dapat dilihat pada bagian 2.5.

1.2 Rumusan Masalah

Dalam tugas akhir ini penulis membahas pembangunan sistem ontologi yang digunakan untuk mengklasifikasikan dokumen tugas akhir mahasiswa jurusan teknik Informatika STT Telkom. Sedangkan modal klasifikasi diambil dari jurnal tugas akhir menggunakan Naïve Bayes Classifier. Sehingga, ontologi tersebut akan terbangun secara otomatis dengan Naïve Bayes Classifier

berdasarkan beberapa jurnal tugas akhir yang sudah diklasifikasikan sebelumnya secara manual.

1.3 Tujuan Tugas Akhir

Dengan tugas akhir ini, diharapkan hal-hal sebagai berikut :

- Membangun perangkat lunak yang dapat mengisi ontologi sebuah Tugas Akhir secara otomatis berdasarkan dokumen jurnal menggunakan Naïve Bayes Classifier.
- Pengujian yang dilakukan adalah untuk melihat tingkat akurasi pengklasifikasian jurnal tugas akhir jika dibandingkan dengan klasifikasi manual. Pengujian tidak akan membandingkan hasil yang didapat oleh metode/algorithm klasifikasi lain.

1.4 Batasan Masalah

Batasan masalah untuk tugas akhir ini adalah sebagai berikut :

- Tugas akhir yang akan dijadikan contoh proses dari sistem ontologi merupakan tugas akhir dari mahasiswa Teknik Informatika STT Telkom.
- Ontologi yang akan dibangun hanya berdasar *class* KBK saja.
- Pengetahuan eksplisit diperoleh dari berkas dokumen jurnal tugas akhir.
- Kakas yang akan digunakan dalam penerapan Naïve Bayes Classifier adalah *Natural Language Toolkit* yang menggunakan metode Orientasi Objek.

1.5 Metodologi

Metode yang digunakan dalam penyelesaian tugas akhir ini adalah :

- Mencari informasi dengan studi pustaka dan referensi dari berbagai sumber seperti artikel, informasi dari buku maupun internet, dan jurnal berkaitan dengan *Knowledge Management, Knowledge Representation, Text Classification, Bayesian Inference*, dan Rekayasa Ontologi.
- Mempelajari konsep ontologi yang akan digunakan pada perangkat lunak.

- Membuat perangkat lunak
- Melakukan pengujian kemampuan perangkat lunak secara fungsional melalui pencarian secara acak terhadap tugas akhir yang telah diarsipkan oleh pihak jurusan.
- Menarik kesimpulan dan menulis laporan.

1.6 **Sistematika Penulisan**

Tugas Akhir ini disusun dengan sistematika pembahasan sebagai berikut :

BAB I PENDAHULUAN

Bab ini memaparkan latar belakang dilakukannya penelitian, perumusan masalah yang akan dibahas, pembatasan masalah, tujuan yang ingin dicapai melalui penelitian ini, metode penyelesaian masalah dan sistematika penulisan.

BAB II DASAR TEORI

Bab ini membahas teori yang mendukung penyusunan tugas akhir ini yaitu mengenai *Knowledge Management*, *Knowledge Representation*, *Text Classification*, *Bayesian Inference*, dan Rekayasa Ontologi.

BAB III ANALISIS DAN PERANCANGAN SISTEM

Menguraikan mengenai proses perancangan sistem dan spesifikasi kebutuhan sistem.

BAB IV IMPLEMENTASI DAN PENGUJIAN

Bab ini menyajikan analisis terhadap representasi ontologi yang digunakan pada sistem dan metode klasifikasi teks menggunakan metode *Naïve Bayes Classifier*.

BAB V PENUTUP

Berisi kesimpulan dari hasil penelitian tugas akhir ini serta saran-saran untuk pengembangan lebih lanjut.

BAB V

PENUTUP

5.1 Kesimpulan

Beberapa kesimpulan yang dapat diambil dari pembangunan Aplikasi NB2Onto ini adalah :

- Naïve bayes layak menjadi alternatif proses klasifikasi teks secara sederhana.
- Proses *training* adalah bagian paling penting dalam proses klasifikasi otomatis, harus selalu diupayakan untuk mengumpulkan data *sample* yang akurat agar proses selanjutnya bisa berjalan dengan benar sesuai dengan yang diharapkan.
- Relatif terhadap Gambar III-1, kontribusi tugas akhir ini hanya sampai pada proses klasifikasi tugas akhir dan belum mulai menyentuh ontologi yang menjadi representasi pengetahuan manajemen tugas akhir jurusan Informatika STT Telkom.
- Keberhasilan proses klasifikasi pada data sampel sebanyak 169 dokumen mencapai tingkat akurasi sebesar 97,4%.

5.2 Saran

Beberapa saran yang bisa diberikan adalah :

- Pihak jurusan Teknik Informatika STT Telkom, dalam hal ini adalah koordinator tugas akhir, sudah harus memikirkan sebuah konsep tentang klasifikasi Tugas Akhir untuk mempermudah pengaturan arsip tugas akhir.
- NLTK adalah kakas yang cukup bagus untuk pengajaran matakuliah pemrosesan bahasa alami, semoga bisa dimanfaatkan kembali untuk riset selanjutnya di bidang linguistik. Karena penulis belum menyentuh beberapa *class* lainnya yang bisa digunakan dalam pemrosesan bahasa

alami. *Class* tersebut antara lain berfungsi untuk membuat visualisasi struktur dan proses-proses pada pemrosesan bahasa alami serta pemrosesan *grammar* bebas konteks dan *finite state automata*.

- Banyak sekali varian dari metoda bayesian yang lebih baik dan dapat digunakan untuk melakukan proses klasifikasi, untuk riset yang lebih mendalam dapat menggunakan bayesian inference atau bayessian network.



DAFTAR PUSTAKA

- [ACL05] The Association for Computational Linguistics. Februari 2005.
<http://www.aclweb.org/>.
- [ANI03] Anindyani R, Dwi. *Pengembangan Aplikasi untuk Pengelolaan Pengetahuan Dokumen Pribadi berbasis Ontologi*. STT Telkom. 2003
- [BOO98] Booch, Grady, James Rumbough, Ivar Jacobson. *The Unified Modeling Language User Guide*. Addison-Wesley Professional; 1st edition. 1998.
- [IND03] Indrajit, Richardus E. *Dasar-dasar Pemrograman Python*. Elex Media Komputindo. 2003.
- [KNR01] *Knowledge Representation*,
<http://www.cirl.uoregon.edu/research/kr.html>. 2001.
- [LOP05] Loper, Edward D & Steven Bird. *Natural Language Toolkit Official Website*. April 2005. <http://nltk.sourceforge.net>.
- [LUG93] Luger dan Stubblefield. *Artificial Intelligence – Structures and Strategies for Complex Problem Solving*. Cummings Publishing Company Inc. 1993.
- [PRE97] Presman, Roger S. *Software Engineering: A Practitioner's Approach*. McGraw-Hill. 1997.
- [ROS05] Rossum, Guido. *Dokumentasi python ver 2.3.4 di linux (/usr/share/docs/python-docs-2.3.4)*. 2005.
- [RUS95] Russel, Stuart J. & Peter Norvig. *Artificial Intelligence: Modern Approach*. Prentice Hall; 1st edition. 1995.
- [SET93] Setiawan, Andi. *Artificial Intelligent*. Andi Offset Yogyakarta. 1993.
- [SMI01] Smith, Barry. *Knowledge Management – Emergency Perspective*.
<http://www.outsight.com/systems/kmgmt/kmgmt.htm>. 2001
- [SUY02] Suyanto. *Intelejensia Buatan*. Bandung: Jurusan Teknik Informatika STT Telkom. 2002.
- [WIK05] Wikimedia Foundation. Mei 2005. <http://www.wikipedia.org>.