

УДК 330.34:340

И.А. Юхно, А.И. Юхно

Харьковский национальный экономический университет, Харьков

ПРИМЕНЕНИЕ ТЕХНОЛОГИЙ ORACLE DATA MINING ПРИ АНАЛИЗЕ ИНТЕРНЕТ-КОНТЕНТА

Современный этап развития экономики Украины требует быстрого реагирования и проведения анализа в различных областях деятельности общества с целью принятия решений подготовки необходимых специалистов. Одной из таких задач является анализ рынка трудоустройства и потребности в специалистах IT-технологий, которая в настоящее время является наиболее динамически развивающейся. Для выполнения поставленной задачи был разработан прототип системы, в которую входили следующие элементы:

- система предварительного сбора информации;
- HTML – страницы из сайта Rabota.ru, Rabota.ua, 101.com.ua;
- пакет Oracle XML[1-3];
- сервер базы данных Oracle Database 11g Enterprise Edition Release 11.0.0.1.0[1];
- программного комплекса с элементами ETL, для помещения данных полученных из Internet в хранилище данных[4,5];
- Oracle Data Miner Server ver.10.2.0.4.1. (Build 543);
- Excel Oracle BI tools. (Build 3.2.3.1.28) [1].

В работе показана возможность проведения сбора, хранения и обработки данных на основе современной СУБД, хранилищ данных и пакетов программ кластерного анализа. В хранилище данных были созданы следующие измерения: специализация; уровень зарплаты; возрастной уровень; географическое расположение работодателей.

Эксплуатация системы в течение нескольких лет позволила сделать следующие выводы:

1. Система показала свою эффективность при анализе потребности в IT-специалистах.
2. Уточнена кластерная модель предметной области.
3. Данная система может применяться для анализа потребностей специалистов других специальностей.
4. Накопление данных в течение длительного времени продемонстрировала временную тенденцию в изменении рынка трудоустройства по различным параметрам.
5. Проведенный анализ результатов может использоваться для планирования и организации обучения по соответствующим профессиональным направлениям.
6. Появление новых версий продуктов расширили возможности системы.

Список литературы

1. [Электронный ресурс]. – Режим доступа к ресурсу: www.oracle.com.
2. Луни К. Oracle database 10g: полный справочник / К. Луни. – М.: Лори, 2006. – Т. 1. – 701 с.
3. Луни К. Oracle database 10g: полный справочник / К. Луни. – М.: Лори, 2006. – Т. 2. – 717 с.
4. Архипенков С. Хранилища данных. От концепции до внедрения / С. Архипенков, Д. Голубев, О. Максименко; под общ. ред. С.Я. Архипенкова. – М.: ДИАЛОГ – МИФИ, 2002 – 528 с.
5. Хоббс Л. Oracle 9i R2: Разработка и эксплуатация хранилищ баз данных / Л. Хоббс, С. Хилсон, Ш. Лоуенд – М.: КУДИЦ – ОБРАЗ, 2004 – 592 с