



Projekt integriert in die Übungen

Einführung in Datenbanksysteme
Datenbanken für die Bioinformatik

Agnès Voisard, Sebastian Müller

1. Anwendungsbeschreibung

Es soll eine Datenbank für Städte und Wetterdaten realisiert werden. Die Datenbank speichert Wettermessungen und Stadtinformationen.

Eine Anwendung stellt Wetterinformationen zu den Städten bereit. Dabei hat nicht jede Stadt eine Wetterstation, aber es soll dennoch möglich sein, das Wetter in dieser Stadt zu ermitteln.

Weiterhin wird die Datenbank für eine Data Mining Anwendung zur Wetterprognose genutzt.

2. Datenherkunft

Die Projektdaten finden sich auf einem Server von [opengeodb.org](http://fa-technik.adfc.de/code/opengeodb/DE.sql). Es wird die folgende Datei benötigt: <http://fa-technik.adfc.de/code/opengeodb/DE.sql> oder <http://fa-technik.adfc.de/code/opengeodb/DE.tab>. Die Wettermessungen können unter <http://dbup2date.uni-bayreuth.de/wetterdaten.html> abgerufen werden.

3. Technologien

Das Projekt wird unter ausschließlicher Verwendung von Java, JDBC und PostgreSQL oder MySQL durchgeführt (für die Web Oberfläche JSP).

Andere relationale Datenbanken sowie die Verwendung von Frameworks wie Hibernate oder Oberflächen wie AWT/Swing sind ausgeschlossen.

4. Teilaufgaben

1. Voraussetzung:

- Installation von PostgreSQL als Server.
- Installation des Java JDK.

2. Anlegen der Datenbank

- Legen Sie eine Datenbank mit dem Namen „stadtweather“ an.

3. Modellierung

- Entwerfen Sie auf Grundlage der Anwendungsbeschreibung und den Daten ein Datenbankschema in, aus der Vorlesung bekannten umgekehrten Chen-Notation mit (min, max) Erweiterung. Hinweis: Die Daten aus opengeodb können weiter normalisiert werden.

4. Übersetzen ins relationale Modell und SQL

- Schreiben Sie die entsprechenden SQL-Queries zur Erstellung der Tabellen. Achten Sie auf eine gute Wahl von Attributeigenschaften wie NOT NULL, UNIQUE und Schlüssel.

5. Datenimport

- Importieren Sie die notwendigen Daten. Definieren Sie Transformationsprozesse zur Umwandlung der importierten Daten in das erstellte Relationenmodell: Definieren Sie einen Transformationsprozess ausschließlich unter Verwendung von SQL Befehlen.
- Definieren Sie einen weiteren Transformationsprozess, in welchem Sie alle Umformungen in Java realisieren und nur durch SELECT und INSERT Statements auf die Datenbank zugreifen (nicht in Kombination :)).

6. Web-Anwendung

- Erstellen Sie eine Web-Anwendung, in welcher der Benutzer zu einer Stadt Wetterinformationen abrufen kann. Hinweis: Die besondere Herausforderung liegt in der Verknüpfung zwischen Wetter- und Stadtinformationen. Hierfür sollte die Position (Längen- und Breitengrad) herangezogen werden, um Wetterstationen in der Nähe finden zu können. Die Berechnung der Distanz zu einer Wetterstation ist nicht trivial: Hilfreich kann das Jcoord Paket sein: <http://www.jstott.me.uk/jcoord/>

7. Data Mining

- Für diese Aufgabe sollen nur die Wetterdaten betrachtet werden.
- Suchen Sie sich eine bestimmte Wetterstation aus (zur Vereinfachung).
- Erstellen Sie ein Feature Set. Es soll eine Aussage über die mittlere Temperatur am Tag x getroffen werden. Mögliche Features sind mittlere Temperatur am Tag x-1 oder Höchsttemperatur der 5 nächsten Wetterstationen am Tag x-1. Sie sollten mindestens 5 Features berechnen können.
- Nutzen Sie das Data Mining Werkzeug weka und wählen Sie einen Klassifikationsalgorithmus, z.B. Naive Bayes. Trainieren Sie Ihren Klassifikator.

5. Weiteres

1. Dokumentation

Dokumentieren Sie im Verlauf des Projektes alle wichtigen Designentscheidungen und den Quellcode, so dass Sie bei der abschließenden Präsentation nachvollziehen können wie der Code funktioniert, was an dieser Stelle geschieht und welche Entscheidungen hier eventuell getroffen wurden. Benutzen Sie hierfür für den Java-Code JavaDoc.

2. Präsentation

Bewertungsgrundlage sind die Zwischenpräsentationen in Anwesenheit eines Tutors.