

Big Data Anwendungen

Datenschutz und gesellschaftliche Aspekte

Vorwort zu diesem Kapitel

- Big Data ist insbesondere deshalb „in aller Munde“...
 - ... weil es neben den technischen Aspekten...
 - ... einige gesellschaftliche Implikationen besitzt
- Bewertung dieser Implikationen stark subjektiv...
 - ... deshalb: kein Abfragen der folgenden Punkte „1:1“ in Klausur...
 - ... aber: Sie sollten sich darüber Gedanken gemacht haben...
 - ... und gegebenenfalls eine fundierte Meinung dazu haben
- Was ich mir für dieses Kapitel wünsche:
 - Ich beziehe bewusst extreme Positionen...
 - ... ich wünsche mir, dass wir „ausgewogen“ diskutieren (nicht nur ich rede)
 - Ich vermute, dass Big Data ändern wird, wie „Märkte“ funktionieren...
 - ... ich wünsche mir, dass Sie das am Ende nachvollziehen können
 - Ich vermute, dass Big Data früher oder später...
 - ... zu grundlegenden, gesellschaftlichen Veränderungen führt...
 - (Ich meine dabei nicht nur, dass Donald Trump Präsident wurde)
 - ... sollte das wirklich eintreten, wünsche ich mir, dass Sie „vorbereitet“ sind

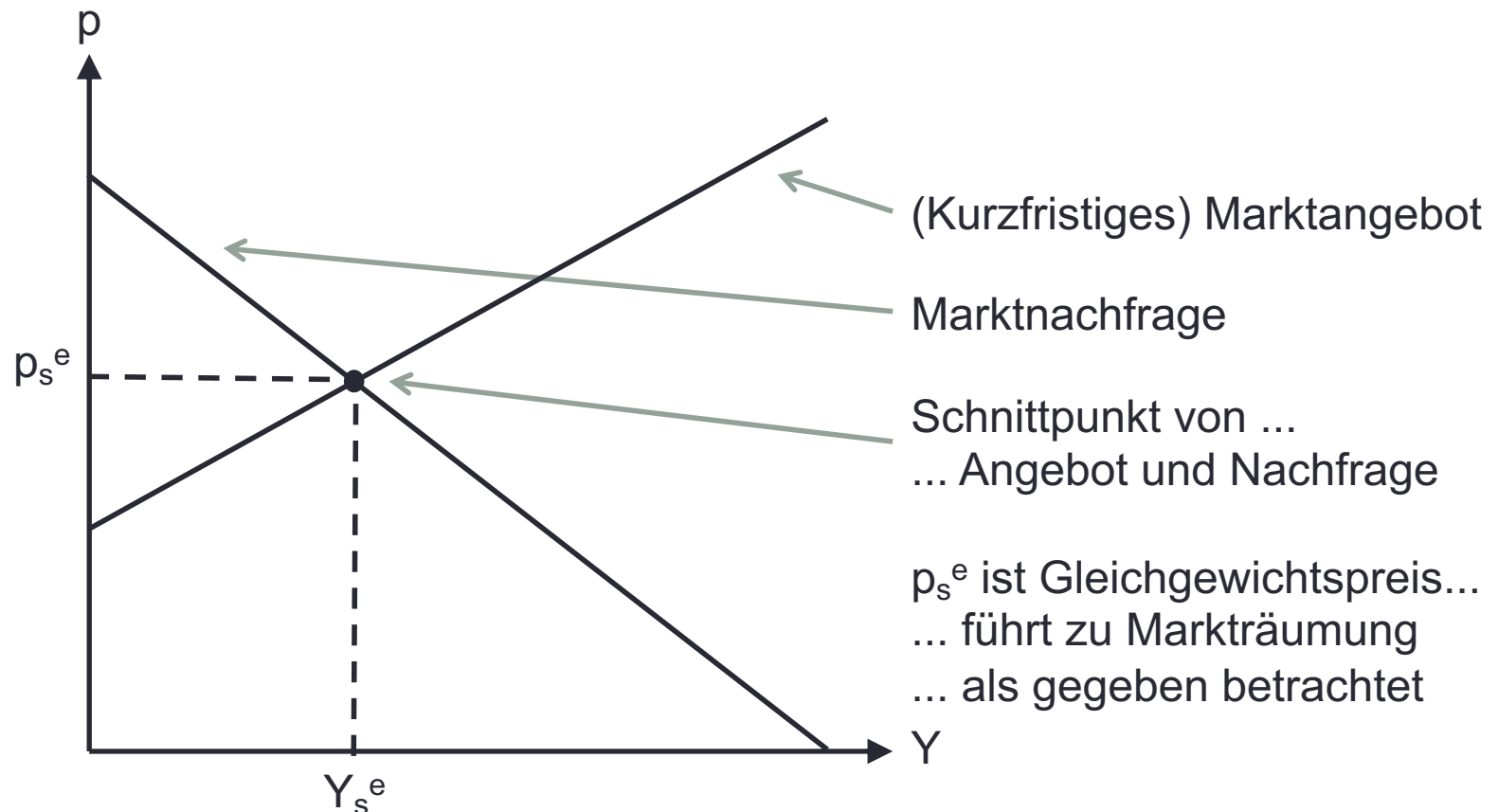
Agenda

- Einführung
- Deskriptive Methoden zur Datenexploration
- Datenqualität
- Klassifikation
- Recommender Systems
- Clusteringverfahren
- Stream Mining
- Social Network Analysis
- Technische Lösungen

- **Datenschutz und gesellschaftliche Aspekte**
 - **Vom Ende des „Marktes“**
 - Von „Jugendsünden“
 - Vom Arbeitsmarkt
 - Zusammenfassung

Der perfekte Markt

- Aus Marktangebot und Marktnachfrage



Der Aktienmarkt

- Überblick
 - Aktienmarkt „ist“ Inbegriff des perfekten Markts
 - Jeder (auch Privatpersonen) kann Kapital anlegen
 - Jeder kann von steigenden bzw. fallenden Kursen profitieren
- Big Data am Aktienmarkt: Hochfrequenzhandel
 - Überschreitet seit 2010 50% Umsatzvolumen des US Aktiengeschäfts
 - Handel läuft hochgradig Algorithmen gestützt (Mensch zu langsam)
 - Mittlere Haltezeit von Aktien ist 22 Sekunden, obwohl...
... Aktien als langfristige Unternehmensbeteiligung gedacht waren
- Strategien
 - Arbitrage: Ausnutzen von Kursdifferenzen an versch. Handelsplätzen
 - Cross Border Handel: Ausnutzen von Kursdifferenzen zwischen Ländern
- Risiken
 - Fehler in Algorithmus führt zu Kettenreaktion...
... Kurse stürzen ab (2010)
 - Privatanleger von der Möglichkeit am Aktienmarkt zu verdienen...
... in Teilen ausgeschlossen

„GAFA Economics“

- Zukunft des „digitalen“ Marktes
 - Wenige Unternehmen (Google, Apple, Facebook und Amazon) besitzen umfassenden „Überblick“ über alle Kundendaten
 - Google/Apple kennen über...
 - ... Smartphones Profile der Kunden
 - ... google mail, docs, ... Meinungen, Termine, Kontakte der Kunden
 - Facebook kennt über...
 - ... Social Network Profile der Kunden
 - ... Loginfunktionalität genutzte Services der Kunden
 - Amazon kennt über...
 - ... Artikelsuchen Konsumverhalten der Kunden
 - ... Amazon Echo wahrscheinlich alles über den Kunden
 - Wenige Unternehmen können die Daten auch zeitnah auswerten
- Wissen gibt Unternehmen Vorteil gegenüber Wettbewerbern...
... sorgt perspektivisch für Nachteile anderer Unternehmen
(GAFA werden perspektivisch Bedarf des Kunden vor ihm selbst kennen)

Agenda

- Einführung
- Deskriptive Methoden zur Datenexploration
- Datenqualität
- Klassifikation
- Recommender Systems
- Clusteringverfahren
- Stream Mining
- Social Network Analysis
- Technische Lösungen

- **Datenschutz und gesellschaftliche Aspekte**
 - Vom Ende des „Marktes“
 - **Von „Jugendsünden“**
 - Vom Arbeitsmarkt
 - Zusammenfassung

Digitales Gedächtnis

- Daten ohne Löschung
 - Mensch erzeugt Daten (unausweichlich)
 - Daten meist durch Dritte (internationale Unternehmen) verarbeitet
 - Datenschutz von Land des Serverstandorts abhängig
 - Datenerzeuger hat (anders als Serverbetreiber) keinen Zugriff auf Daten
- PRISM
(Planning Tool for Resource Integration, Synchronization, and Management)
 - Programm der NSA zur Auswertung elektronischer Daten
(unter Beteiligung von Microsoft, Google, Facebook, Yahoo, Apple, AOL)
 - Als rechtschaffener Bürger haben Sie hoffentlich nichts zu verbergen
 - Selbst wenn das Projekt jetzt beendet würde, hätte der NSA über Jahre...
... hinweg die Überwachungsdaten zur Bevölkerung (hätte die Stasi ...)
- Data Leaks
 - Sicherstellung von Datensicherheit schwierig / teuer
 - „Verlust“ von Daten verhältnismäßig günstig
 - Anreiz Daten zu schützen oft überschaubar

Eigentum von Daten

- Mögliche Eigentümer von Daten
 - Datenpflegender Benutzer
 - Individualperson zuordnender Benutzer
 - Datenerfassendes Unternehmen
 - Datenspeicherndes Unternehmen
 - Unternehmen dessen AGB angenommen wurden
- Mögliche Umgangsrechte mit Daten
 - Verknüpfen der Daten mit anderen Daten
 - Verkauf der Daten
 - Weitergabe der Daten
- Durchsetzung von Eigentumsrechten
 - Internationale Umsetzung problematisch
 - Internationale Institutionen zur Rechtsprechung unklar

Recht auf informationelle Selbstbestimmung

- Deutsche Staatsbürger haben „Recht auf informationelle Selbstbestimmung“
 - Bürger kann über Verwendung personenbezogener Daten bestimmen
 - Personenbezogene Daten sind individueller Person direkt zuordenbar
 - Internationale Durchsetzung des Rechts problematisch
 - Oft Datenaufkommen in Unternehmen selbst unklar
(habe selbst Auskunftersuchen gestellt – Ergebnis ist unbefriedigend)
- Bester Schutz: Eigenverantwortlicher Umgang mit Daten
 - Vor Datenerhebung prüfen wer Daten erhält
 - Daten verschiedenen Anbietern anvertrauen
 - Wo möglich deutsche oder europäische Anbieter Nutzen
 - Prüfen ob Mehrwert die Weitergabe entsprechender Daten rechtfertigt
 - Daten sparsam weitergeben (galt übrigens schon vor dem Internet)
 - Nicht auf den guten Willen des Datenhalters vertrauen
(Im Zweifel wird immer gemacht, was juristisch möglich ist)

Agenda

- Einführung
- Deskriptive Methoden zur Datenexploration
- Datenqualität
- Klassifikation
- Recommender Systems
- Clusteringverfahren
- Stream Mining
- Social Network Analysis
- Technische Lösungen

- **Datenschutz und gesellschaftliche Aspekte**
 - Vom Ende des „Marktes“
 - Von „Jugendsünden“
 - **Vom Arbeitsmarkt**
 - Zusammenfassung

Konsequenzen für den Arbeitsmarkt

- Big Data kann menschliche Tätigkeiten automatisieren
 - Automatische Beantwortung von Kunden-E-Mails
 - Qualität meist katastrophal (wird aber besser)
 - Stellen im Bereich Kundensupport entfallen
 - Automatische Disposition von Ware
 - Qualität der Vorhersage oft besser als „Bauchgefühl“ des Disponenten
 - Stellen im Bereich Disposition entfallen
 - Angebotsgestaltung für Kunden
 - Qualität automatischer Prognose oft besser als individuelle Beratung
 - Stellen im Bereich Verkauf entfallen
 - ...
- Gefahr Menschen durch Algorithmen zu ersetzen
 - Zunehmend mehr Arbeitsplätze im Bereich Algorithmenentwicklung
 - Gefahr des Ersetzens „traditioneller“ Arbeitsplätze

Agenda

- Einführung
- Deskriptive Methoden zur Datenexploration
- Datenqualität
- Klassifikation
- Recommender Systems
- Clusteringverfahren
- Stream Mining
- Social Network Analysis
- Technische Lösungen

- **Datenschutz und gesellschaftliche Aspekte**
 - Vom Ende des „Marktes“
 - Von „Jugendsünden“
 - Vom Arbeitsmarkt
 - **Zusammenfassung**

Zusammenfassung

- Wahrscheinlich sind Gefahren überschaubar
 - Auswirkung auf den Markt
Sollte hier tatsächlich ein Ungleichgewicht eintreten...
... wird staatliche Regulierung dem entgegenwirken
 - Datengedächtnis
Erfahrungsgemäß sind junge Menschen heute schon...
... deutlich weniger sorglos als ältere, da sie Zusammenhänge verstehen
 - Arbeitsmarkt
Potentiell irrelevant, da jede gesellschaftliche Änderung...
... entsprechende Änderungen am Arbeitsmarkt mit sich bringt...
... sollte am Ende weniger Arbeit für alle übrigbleiben ist das OK
- Trotzdem
 - Wichtig sich die Auswirkungen bewusst zu machen
 - Wichtig aufmerksam zu sein
 - Wichtig eigenes Verhalten gegebenenfalls anzupassen