

Projektplan Automatidata

Projektname: Vorhersage von Taxitarifen in New York

Projektleiter: Dominik Vogel

Datenwissenschaftler: Dominik Vogel

Business Intelligence Analyst: Dominik Vogel

Data Engineer: Dominik Vogel

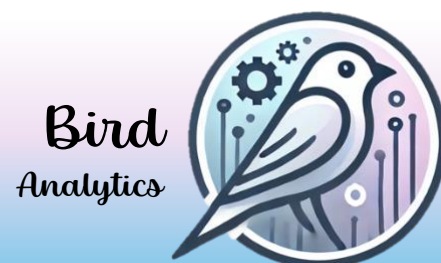
Chief Data Officer: Dominik Vogel

Projektstartdatum: [Startdatum]

Projektenddatum: [Enddatum]

Zusammenfassung des Projekts:

Dieses Projekt zielt darauf ab, ein präzises Vorhersagemodell für Taxitarife in New York City zu entwickeln. Mithilfe der Daten der New York City Taxi und Limousinen Kommission sollen Tarifstrukturen optimiert und den Fahrgästen bessere Preisvorhersagen ermöglicht werden. Die Hauptphasen umfassen Datenanalyse, Modellentwicklung, Validierung und Präsentation der Ergebnisse.



Inhaltsangabe

1. **Projektübersicht**
2. **Projektstruktur nach PACE**
3. **Meilensteine und Zeitrahmen**
4. **Risiken und Abhängigkeiten**
5. **Nächste Schritte**
- 6.



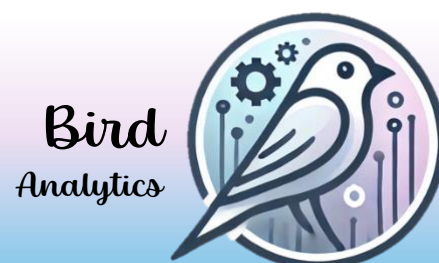
Projektplan nach PACE

1. Projektübersicht

- **Projektname:** Vorhersage von Taxitarifen in New York City
- **Projektziel:** Entwicklung eines präzisen Modells zur Vorhersage von Taxitarifen basierend auf den gesammelten Daten der New York City Taxi and Limousine Commission. Dies soll dazu beitragen, Tarifstrukturen zu optimieren und den Fahrgästen eine bessere Preisvorhersage zu ermöglichen.
- **Projektumfang:**
 - **Inklusive:** Datenanalyse, Modellentwicklung, Validierung, Ergebnispräsentation
 - **Exklusive:** Langfristige Modellwartung, Integration in operative Systeme

- **Stakeholder:**

Rolle	Name	Verantwortung
Projektleiter (PL)	Dominik Vogel	Koordination & Kontrolle
Data Scientist (DS)	Dominik Vogel	Datenanalyse & Modellbau
Business Intelligence (BI)	Dominik Vogel	Reporting & Visualisierung
Data Engineer (DE)	Dominik Vogel	Datenzugang & Vorbereitung
Chief Data Officer (CDO)	Dominik Vogel	Data Governance



2. Projektstruktur nach PACE

PACE-Phase	Aufgabe	Deliverables	Verantwortlicher	Dauer
Plan	Projektstrategie erstellen	Strategie-Dokument	Projektleiter	1-2 Tage
	Projektvorschlag erstellen	Projektvorschlag	Data Scientist	1-2 Tage
	Datenquellen und Tools identifizieren	Liste der Datenquellen/Tools	Data Engineer	1-2 Tage
	Stakeholder-Briefing	Stakeholder-Kommunikation	Projektleiter	1 Tag
Analyze	Datenzugriff und -vorbereitung	Rohdatensätze	Data Engineer	2-3 Tage
	Explorative Datenanalyse (EDA)	EDA-Bericht	Data Scientist	1 Woche
	Datenbereinigung und Transformation	Bereinigte Datensätze	Data Scientist	1-2 Wochen
	Hypothesenbildung und Deskriptive Analysen	Hypothesentest-Ergebnisse	Data Scientist	1 Woche
Construct	Modell auswählen und erstellen	Prototyp-Modell	Data Scientist	2 Wochen
	Modelltraining und -optimierung	Optimiertes Modell	Data Scientist	2 Wochen
	Modellvalidierung	Validierungsbericht	Data Scientist	1 Woche
	Visualisierungen entwickeln	Dashboards und Graphiken	BI Analyst	1 Woche
Execute	Ergebnisse präsentieren	Präsentation/Executive Summary	Projektleiter/BI Analyst	1 Woche
	Feedback von Stakeholdern einholen	Feedback-Dokument	Projektleiter	2-3 Tage
	Modell überarbeiten	Finalisiertes Modell	Data Scientist	1 Woche
	Abschlussbericht erstellen	Endgültiger Projektbericht	Projektleiter	1 Woche

3. Meilensteine und Zeitrahmen

1. **Meilenstein 1:** Projektinitiierung (1 Woche)

- Aufgaben:
 - Strategie-Dokument erstellen
 - Stakeholder identifizieren und briefen
 - Projektvorschlag finalisieren
- Deliverables: Strategie-Dokument, Projektvorschlag, Stakeholder-Liste

2. **Meilenstein 2:** Datenanalyse und Exploration (2-3 Wochen)

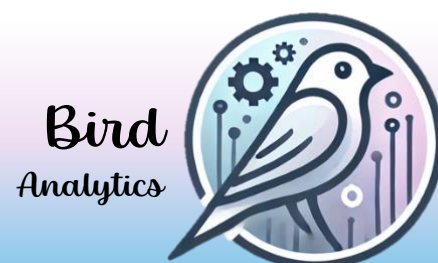
- Aufgaben:
 - Datenzugang sicherstellen und Datenqualität prüfen
 - Explorative Datenanalyse (EDA) durchführen
 - Daten bereinigen und vorbereiten
- **Deliverables:** EDA-Bericht, bereinigte Datensätze

3. **Meilenstein 3:** Modellentwicklung und Validierung (3-4 Wochen)

- Aufgaben:
 - Regressions- oder ML-Modell erstellen
 - Modell trainieren und optimieren (Hyperparameter-Tuning)
 - Modell validieren und Testergebnisse analysieren
- **Deliverables:** Validierungsbericht, optimiertes Modell

4. **Meilenstein 4:** Präsentation und Feedback (2 Wochen)

- Aufgaben:
 - Ergebnisse visualisieren und präsentieren
 - Stakeholder-Feedback einarbeiten
 - Abschlussbericht erstellen
- **Deliverables:** Präsentation, Feedback-Dokument, finaler Bericht



4. Risiken und Abhängigkeiten

Risiko	Auswirkung	Maßnahme
Fehlender Datenzugriff	Verzögerung des Projektstarts	Frühzeitige Absprache mit IT/DE
Datenqualität unzureichend	Ungenaue Modelle/Ergebnisse	Datenbereinigung und QA durchführen
Modell erfüllt Genauigkeitsziel nicht	Ergebnisakzeptanz gefährdet	Alternative Modelle/Optimierungen
Ressourcenknappheit	Verzögerte Fertigstellung	Aufgaben priorisieren, Team erweitern
Feedback der Stakeholder verzögert	Endgültige Freigabe wird verzögert	Regelmäßige Updates und Follow-ups



5. Nächste Schritte

1. Strategieplanung und Kickoff-Meeting

- Review der Strategie-Dokumente und Projektvorschläge
- Klärung der Rollenverteilung (RACI-Matrix)
- Kickoff-Meeting mit Stakeholdern durchführen

2. Datenzugang sicherstellen

- Klärung der Zugriffsrechte für alle Datenquellen
- Validierung der Datenqualität

3. Datenanalyse-Phase starten

- Explorative Datenanalyse und Hypothesentests durchführen
- Bereinigung und Transformation der Daten

4. Modellentwicklungs-Phase

- Modellarchitektur auswählen (z.B. ML-Algorithmen wie Regression, XGBoost)
- Training und Optimierung der Modelle durchführen

5. Ergebnisse dokumentieren und präsentieren

- Erstellung von Visualisierungen (Tableau, Power BI)
- Ergebnispräsentation für Stakeholder vorbereiten

6. Feedbackrunde und Überarbeitung

- Stakeholder-Feedback einholen und einarbeiten
- Endgültiges Modell und Bericht finalisieren

