

PRODUCT REQUIREMENTS DOCUMENT (PRD)

Självförbättrande Analyssystem för Dagens Dubbel – Solvalla Lördagstrav

Version: 1.0

Datum: 7 februari 2026

Författare: Kent Lundgren

Status: Initial specifikation

1. SAMMANFATTNING

1.1 Produktbeskrivning

Ett självförbättrande analyssystem som predikterar vinnare i Dagens Dubbel på Solvallas lördagstrav. Systemet analyserar travlopp baserat på multipla faktorer (ej enbart odds), utvärderar sina förutsägelser mot faktiska resultat, och justerar kontinuerligt sina viktningsparametrar för att förbättra träffssäkerheten över tid.

1.2 Affärsmål

- Skapa ett datadrivet beslutsstöd för travspel
- Uppnå progressivt ökad träffssäkerhet genom maskininlärning
- Dokumentera vilka faktorer som är mest prediktiva för travlopp
- Generera kvantitativ ROI-analys över tid

1.3 Målgrupp

- Primär: Användaren själv (Kent Lundgren)
 - Sekundär: Potentiella framtida användare med intresse för datadriven travanalys
-

2. DAGENS DUBBEL – KONTEXT OCH SPELFORM

2.1 Spelformens struktur

Dagens Dubbel (DD) består av de två sista loppen i V85-omgången på Solvalla lördagar. För att vinna måste man pricka vinnaren i båda dessa lopp.

Exempel från idag (7 februari 2026):

- V85 består av 8 lopp
- Dagens Dubbel = Lopp 7 + Lopp 8
- Starttid: ca 14:35 (första loppet)
- Kontext: STL-finaler och Margareta Tidiga Unghästserie

2.2 Utmaningen

Dagens Dubbel kräver att träffa rätt i båda loppen, vilket gör analysen särskilt viktig. Odds återspeglar publikens kollektiva bedömning, men systematisk analys av objektiva faktorer kan identifiera undervärderingar.

3. ANALYSEFAKTORER

3.1 Primära analysfaktorer (MÅSTE analyseras)

3.1.1 Hästarnas prestationshistorik

- **Senaste 5 loppen:** Placering, tid, margin
- **Formkurva:** Trend uppåt/nedåt/stabil
- **Vinst-procent:** Totalt och senaste 3 månaderna
- **Placeringsprocent:** Top 3-placeringar

3.1.2 Startspår och position

- **Startnummer:** Inre vs yttre spår
- **Statistik per spår:** Historisk vinstfrekvens från varje spår
- **Spårets betydelse per distans:** Kort vs lång distans

3.1.3 Kusk och tränare

- **Kuskens statistik:** Vinstprocent, placeringsprocent
- **Tränarens statistik:** Formkurva, specialisering
- **Kombo-statistik:** Hur kusk + tränare presterar tillsammans
- **Hemmapanevördel:** Om det är Solvalla-tränare

3.1.4 Distans och startmetod

- **Hästen på distansen:** Bästa tid, antal starter, vinstprocent
- **Startmetod:** Autostart vs voltstart
- **Distansförändring:** Hur hästen presterat vid förra distansändring

3.1.5 Tävlingspaus

- **Dagar sedan senaste loppet**
- **Prestationsmönster efter paus:** Bättre/sämre efter vila?
- **Antal starter senaste 3 månaderna**

3.1.6 Loppets karaktär

- **Propositionsklass:** Vilken nivå hästen normalt tävlar i
- **Prispengar:** Loppets status
- **Finaler vs försökslopp:** STL-final vs vanligt lopp
- **Unghäst vs erfaren:** Margareta-serie vs öppna lopp

3.2 Sekundära analysfaktorer (BÖR analyseras)

3.2.1 Hästarnas fysiska data

- **Ålder:** Toppålder för travare (4-8 år typiskt)
- **Kön:** Hingst/valack/sto-statistik
- **Stamtavla:** Framgångsrika släktlinjer

3.2.2 Tillbehör och utrustning

- **Barfota vs skor:** Ändring i balanseringstillbehör
- **Skoddsändringar:** Indikerar tränarens förväntningar
- **Utrustningsändringar:** Skyggapp, seletöy etc.

3.2.3 Kontextuella faktorer

- **Väder och underlag:** Temperatur, nederbörd, banans status
- **Tidpunkt på dagen:** Ljusförhållanden, banastighet
- **Antal starter i loppet:** Fältstorlek

3.2.4 Historisk matchup

- **Head-to-head:** Hur hästar presterat mot varandra
- **Rivalitet:** Kända dueller

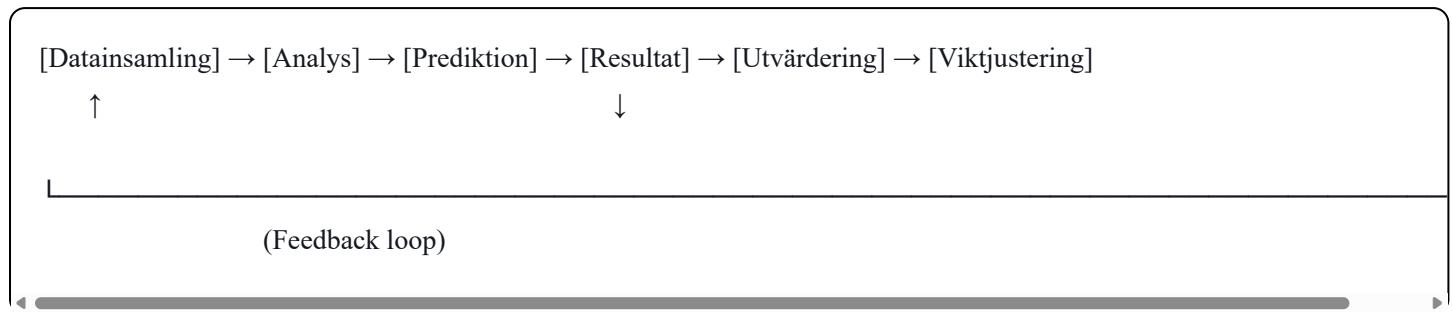
3.3 Odds (Referensfaktor, EJ primär)

- **Vinnarodds:** Publikens uppfattning
- **Oddstrend:** Förändring från morgon till start
- **Oväntade favoriter:** Oddsskillnad mellan expertis och publik

VIKTIGT: Odds ska dokumenteras men inte vara primär grund för analys. Målet är att hitta undervärderingar genom faktaanalys.

4. SYSTEMARKITEKTUR

4.1 Dataflöde



4.2 Komponenter

4.2.1 Datainsamling

Datakällor:

- ATG.se (startlistor, resultat, statistik)
- Svensk Travsport (propositioner, hästar, kuskar)
- Travsport.se (historisk data)
- SMHI (väderdata)

Datapunkter att samla (per häst, per lopp):

- Hästidentitet (namn, ID, ålder, kön, stamtavl)
- Startposition (spår, startnummer)
- Kusk och tränare (namn, ID, hemmatrännare)
- Historiska resultat (senaste 10 loppen minimum)
- Loppets egenskaper (distans, startmetod, klass, prispengar)
- Odds (vid flera tidpunkter: morgon, 1h före, start)
- Resultat (placering, tid, margin, galopp/diskvalifikation)

4.2.2 Analysmotor

Viktningsmodell (initial):

Faktor	Initial vikt	Rationale
Formkurva (senaste 5)	25%	Aktuell form är högst relevant
Startnummer/spår	15%	Stor betydelse på kort distans
Kusk-statistik	15%	Professionell kompetens avgörande
Distans-lämplighet	15%	Hästar har specialdistanser
Tränare-statistik	10%	Långsiktig kompetens
Tävlingspaus	8%	För lång/kort paus påverkar
Loppets klass	7%	Kliv upp/ned i klass viktig faktor
Head-to-head	5%	Mindre vikt, begränsad data

VIKTIGT: Dessa vikter är utgångspunkter och ska justeras baserat på utvärdering.

4.2.3 Prediktionsmotor

För varje häst i DD-loppen:

1. **Poängberäkning:** Sumdera viktade faktorer till totalscore
2. **Rankning:** Rangordna hästar 1-12 (eller antal startande)
3. **Sannolikhetsfördelning:** Konvertera poäng till vinstsannolikhet
4. **DD-kombinationer:** Generera top-kombinationer (t.ex. top 3 × top 3)

Output:

- Rankning per lopp
- Rekommenderade DD-kombinationer
- Konfidensintervall

4.2.4 Utvärderingsmotor

Efter varje lördag:

1. **Resultatinsamling:** Dokumentera faktiska vinnare + fullständigt resultat
2. **Träffanalys:**
 - Träffade systemet rätt (DD-kombination)?
 - Placering av systemets topphästar (top 1, top 3, top 5)?
 - Vilka faktorer kännetecknade vinnarna?
3. **Faktorsignifikans:**

- Korrelation mellan faktorer och seger
- Vilka vikter gav för hög/låg betydelse?

4. Oddsjämförelse:

- Var vinnarna övervärderade/undervärderade i odds?
- Lyckades systemet identifiera value bets?

4.2.5 Viktjusteringsmotor

Baserat på utvärdering:

1. Faktorjustering:

- Öka vikt för faktorer som korrelerade med seger
- Minska vikt för faktorer med svag korrelation

2. Spårjustering:

- Uppdatera spårstatistik baserat på faktiska resultat

3. Kusk/tränare-justering:

- Dynamisk uppdatering av statistik

4. Algoritm-finslipning:

- Identifiera icke-linjära samband
- Lägg till interaktionseffekter (t.ex. kusk+häst-synergi)

5. DATAINSAMLING OCH STRUKTUR

5.1 Pre-race data (samlas före lopp)

Filformat: JSON eller CSV

json

```
{  
  "race_id": "solvalla_2026-02-07_DD1",  
  "race_number": 7,  
  "date": "2026-02-07",  
  "track": "Solvalla",  
  "distance": 2140,  
  "start_method": "auto",  
  "race_class": "STL-final",  
  "prize_money": 300000,  
  "weather": {  
    "temperature": -2,  
    "conditions": "clear",  
    "track_condition": "fast"  
  },  
  "horses": [  
    {  
      "post_position": 1,  
      "name": "Example Horse",  
      "age": 5,  
      "sex": "gelding",  
      "driver": "Driver Name",  
      "trainer": "Trainer Name",  
      "home_trainer": true,  
      "last_5_races": [  
        {"date": "2026-01-31", "position": 2, "time": "1:11.5", "margin": 0.5},  
        {"date": "2026-01-24", "position": 1, "time": "1:11.2", "margin": 0},  
        ...  
      ],  
      "career_stats": {  
        "starts": 45,  
        "wins": 12,  
        "places": 25,  
        "earnings": 850000  
      },  
      "distance_stats": {  
        "2140m": {"starts": 15, "wins": 5, "best_time": "1:10.8"}  
      },  
      "days_since_last_race": 7,  
      "odds": {  
        "morning": 4.5,  
        "1h_before": 4.2,  
        "start": 3.8  
      }  
    },  
    ...
```

```
]
}
```

5.2 Post-race data (samlas efter lopp)

```
json

{
  "race_id": "solvalla_2026-02-07_DD1",
  "result": [
    {"position": 1, "post_position": 5, "name": "Winner Horse", "time": "1:10.5", "odds": 3.2},
    {"position": 2, "post_position": 2, "name": "Second Horse", "time": "1:10.7", "margin": 0.2},
    ...
  ],
  "dd_payout": 156.50,
  "system_prediction": {
    "predicted_winner_race1": 5,
    "predicted_winner_race2": 3,
    "actual_winner_race1": 5,
    "actual_winner_race2": 7,
    "hit_race1": true,
    "hit_race2": false,
    "dd_hit": false
  }
}
```

5.3 Evaluation data (genereras efter analys)

```
json
```

```
{
  "week": "2026-W06",
  "date": "2026-02-07",
  "accuracy": {
    "dd_hit": false,
    "race1_top3": true,
    "race2_top3": false,
    "race1_winner_rank": 1,
    "race2_winner_rank": 5
  },
  "factor_performance": {
    "form_curve": {"correlation": 0.72, "weight_adjustment": +2},
    "post_position": {"correlation": 0.45, "weight_adjustment": -3},
    "driver_stats": {"correlation": 0.68, "weight_adjustment": +1},
    ...
  },
  "updated_weights": {
    "form_curve": 27,
    "post_position": 12,
    "driver_stats": 16,
    ...
  }
}
```

6. UTVÄRDERINGSMETODIK

6.1 Successmått (KPIs)

6.1.1 Primära KPIs

- **DD-träffprocent:** Antal korrekta DD-förutsägelser / totalt antal veckor
 - Mål vecka 1-4: >5%
 - Mål vecka 5-12: >10%
 - Mål vecka 13+: >15%
- **ROI (Return on Investment):** (Totala vinster - Totala insatser) / Totala insatser
 - Mål långsiktig: >0% (breakeven)
 - Stretch goal: >20%
- **Top-3 träffssäkerhet:** Andel gånger vinnaren rankades i systemets top 3
 - Mål vecka 1-4: >40%
 - Mål vecka 5+: >60%

6.1.2 Sekundära KPIs

- **Faktorkorrelation:** R^2 mellan faktorer och seger
- **Oddsifferens:** Genomsnittlig avvikelse mellan systemodd och faktiskt odd
- **Vinsttid-trend:** Förbättring i viktningsmodell över tid
- **False positive rate:** Hur ofta top-rankade hästar floppade

6.2 Utvärderingsfrekvens

- **Veckovis:** Efter varje lördag, grundläggande utvärdering
- **Månadsvis:** Djupare analys av trender och justeringsbehov
- **Kvartalsvis:** Omfattande revidering av algoritm och faktorer

6.3 A/B-testning av algoritmer

För att säkerställa verkliga förbättringar:

- Spara "frysta" versioner av algoritmen
 - Jämför ny version med gammal version retroaktivt
 - Implementera endast förbättringar som är statistiskt signifikanta ($p < 0.05$)
-

7. TEKNISK IMPLEMENTATION

7.1 Tech stack (förslag)

Backend:

- **Språk:** Python 3.11+
- **Data processing:** Pandas, NumPy
- **Machine learning:** Scikit-learn (initial), ev. TensorFlow/PyTorch senare
- **Web scraping:** BeautifulSoup, Selenium (för ATG-data)
- **API:** FastAPI (om webbgränssnitt önskas)

Data storage:

- **Databas:** SQLite (initial), PostgreSQL (skalning)
- **Fillagring:** JSON/CSV för backup och export

Frontend (valfritt):

- **HTML/CSS/JavaScript:** För visualisering
- **Chart.js / Plotly:** För grafer och statistik

7.2 Modulstruktur

```
travanalys/
    ├── data_collection/
    |   ├── scraper.py      # Hämta data från ATG, Transport
    |   ├── weather.py     # Hämta väderdata
    |   └── database.py    # Databasoperationer
    ├── analysis/
    |   ├── feature_engine.py # Beräkna analysfaktorer
    |   ├── weighting.py    # Viktningsmodell
    |   └── prediction.py   # Prediktionsmotor
    ├── evaluation/
    |   ├── results.py      # Samla resultat
    |   ├── performance.py  # Beräkna KPIs
    |   └── adjustment.py   # Viktjustering
    ├── utils/
    |   ├── config.py       # Konfiguration
    |   └── logger.py       # Loggning
    ├── output/
    |   ├── reports.py      # Generera rapporter
    |   └── visualizations.py # Skapa grafer
    └── main.py            # Huvudprogram
```

7.3 Workflow

Fredag kväll / Lördag morgon:

1. Scrape startlistor för dagens DD
2. Samla historisk data för alla hästar
3. Beräkna analysfaktorer
4. Generera prediktioner
5. Output rapport med rekommendationer

Lördag eftermiddag (efter lopp): 6. Samla resultat 7. Utvärdera prestanda 8. Justera vikter 9. Spara data till databas 10. Generera utvärderingsrapport

8. OUTPUT OCH RAPPORTERING

8.1 Pre-race rapport

Format: HTML eller PDF

Innehåll:

- Sammanfattning av dagens DD-lopp

- Rankning av hästar per lopp (med poäng)
- Rekommenderade DD-kombinationer (top 5)
- Faktor-breakdown för topphästar
- Oddsjämförelse
- Konfidensintervall

Exempel:

==== DAGENS DUBBEL – SOLVALLA 7 FEBRUARI 2026 ===

LOPP 7 – STL-final, 2140m autostart

-
1. Example Horse (spår 5) – Score: 8.7/10
 - Formkurva: Excellent (4 segrar senaste 5)
 - Kusk: Top-class (22% vinstdel)
 - Spår: Optimalt för distansen
 - Odds: 3.8 (UNDERVÄRDERAD)

2. Second Horse (spår 2) – Score: 7.9/10
- ...

LOPP 8 – Margareta-serie, 2140m autostart

-
1. Young Star (spår 3) – Score: 8.2/10
- ...

REKOMMENDERADE DD-KOMBINATIONER:

1. Example Horse + Young Star (sannolikhet: 18%)
 2. Example Horse + Challenger (sannolikhet: 14%)
 3. Second Horse + Young Star (sannolikhet: 12%)
- ...

VIKTNINGSMODELL DENNA VECKA:

- Formkurva: 25%
 - Startnummer: 15%
 - Kusk: 15%
- ...

8.2 Post-race utvärdering

Format: HTML eller PDF

Innehåll:

- Faktiska resultat
- Systemets träffsäkerhet

- Faktorenanalys (vilka faktorer kännetecknade vinnaren?)
- Viktjusteringar
- Kumulativ prestanda (ROI, träffprocent)
- Grafer och trendanalys

Exempel:

==== UTVÄRDERING – SOLVALLA 7 FEBRUARI 2026 ===

RESULTAT:

Lopp 7: Vinnare = Example Horse (systemet rankade #1) ✓

Lopp 8: Vinnare = Underdog (systemet rankade #5) X

DD-TRÄFF: NEJ

FAKTORER SOM KÄNNETECKNADE VINNARNA:

Lopp 7:

- Stark formkurva (4/5 vinster)
 - Inre spår
 - Erfaren kusk
- Systemet identifierade rätt!

Lopp 8:

- Svag formkurva (0/5 vinster) – oväntat
 - Yttre spår (#11)
 - Ny tränare (specialiserad på unghästar)
- Systemet missade: Vikt på "nytt tränarsamarbete" var för låg

VIKTJUSTERINGAR:

- "Ny tränare (unghäst)" +3% (5% → 8%)
- "Spår (långdistans)" -2% (15% → 13%)

KUMULATIV PRESTANDA (6 veckor):

- DD-träffar: 1/6 (16.7%)
- Top-3 träffssäkerhet: 9/12 (75%)
- ROI: -12% (förbättring från -25% vecka 1)

9. ITERATIV FÖRBÄTTRING

9.1 Feedback-loop

Varje veckas utvärdering matar systematiskt in i nästa veckas analys:

1. **Vikter justeras:** Baserat på faktorkorrelation
2. **Nya faktorer introduceras:** Om mönster upptäcks (t.ex. "ny tränare för unghäst")

3. **Outliers analyseras:** Extrema resultat studeras för nya insikter
4. **Modellkomplexitet ökar:** Från linjär viktning till icke-linjära samband

9.2 Hypotesgenerering

Systemet ska aktivt generera hypoteser att testa:

Exempel:

- *Hypotes 1:* "Hästar från Djuse-stallet vinner oftare på Solvalla-lördagar (hemmabanevördel)"
 - Test: Jämför Djuse-hästar vs andra tränare över 12 veckor
 - Resultat: Om signifikant → lägg till "hemmatrännare-bonus" i viktning
- *Hypotes 2:* "Kuskar med >20% vinstprocent ger extra value vid odds >5.0"
 - Test: Analysera korrelation mellan topkuskar och höga odds
 - Resultat: Om sant → justera modell för value bets

9.3 Månadsrapport

Varje månad: omfattande analys med följande delar:

1. Prestanda-dashboard:

- DD-träffar denna månad
- ROI-utveckling (graf)
- Faktorbetydelse (heatmap)

2. Upptäckta mönster:

- Vilka faktorkombon fungerar bäst?
- Vilka spår är övervärderade/undervärderade?
- Vilka kuskar/tränare är hot/cold?

3. Modelljusteringar:

- Vad har ändrats i viktning?
- Nya faktorer som lagts till
- Borttagna faktorer (för låg korrelation)

4. Nästa månads fokus:

- Vilka hypoteser ska testas?
- Vilka datapunkter saknas?
- Planerad utveckling

10. RISKER OCH BEGRÄNSNINGAR

10.1 Datarisker

- **Datakvalitet:** Felaktig eller ofullständig data från källor
- **Scraping-begränsningar:** ATG kan ändra webbplatsstruktur
- **Historisk bias:** Gamla data kanske inte längre relevanta

Mitigering:

- Validera data mot flera källor
- Bygga robust scraper med felhantering
- Viktad tidsavtagning (nyare data viktigare)

10.2 Modelrisker

- **Overfitting:** Modellen anpassas för mycket till historisk data
- **Små sample size:** Begränsat antal lopp kan ge slumpmässiga resultat
- **Black swan events:** Oväntade händelser (häst sjuk, väderskifte, startbilsfel)

Mitigering:

- Korsvalidering
- Behåll modellens enkelhet initialt
- Dokumentera outliers separat

10.3 Användningsrisker

- **Spelberoende:** Systemet är verktyg, inte garanti för vinst
- **Ekonomisk risk:** Spel innehåller alltid risk för förlust

Mitigering:

- Tydlig disclaimer: "Endast beslutsstöd, aldrig garanti"
- Rekommenderad max insats per vecka
- Tracking av totalbalans (förluster inkluderade)

11. TIDSPLAN OCH MILSTOLPAR

Fas 1: MVP (Vecka 1-4)

Mål: Grundläggande fungerande system

- Vecka 1: Datainsamling och struktur
- Vecka 2: Implementera grundläggande viktningsmodell
- Vecka 3: Generera första prediktionen
- Vecka 4: Samla första resultat och utvärdera

Leverans: Enkel HTML-rapport med prediktioner och utvärdering

Fas 2: Förbättring (Vecka 5-12)

Mål: Kontinuerlig förbättring av modell

- Justering av vikter varje vecka
- Introducera nya faktorer baserat på upptäckter
- A/B-testning av algoritmer
- Bygga historisk databas

Leverans: Månadsrapporter med trendanalys

Fas 3: Optimering (Vecka 13+)

Mål: Maximera träffsäkerhet och ROI

- Införa maskininlärning (om datamängd tillräcklig)
- Avancerade statistiska metoder
- Automatiserad hypotesgenerering
- Prediktiv modellering

Leverans: Fullt automatiserat system med dashboards

12. FRAMTIDA UTVECKLING

12.1 Kortsiktig (3-6 månader)

- Utöka till andra travbanor (Jägersro, Åby, etc.)
- Inkludera fler spelformer (V75, V86)
- Bygga webbgränssnitt för enkel åtkomst

12.2 Långsiktig (6-12 månader)

- Machine learning-modeller (neural networks, random forests)
- Realtidsanalys under lopp (om live-data tillgänglig)
- Automatisk spelplacering (via API)

- Community-funktioner (dela prediktioner, jämföra resultat)

12.3 Blue sky (12+ månader)

- Integration med video-analys (hästarnas gångstil, kondition)
 - Sentiment-analys från expertkommentarer
 - Ensemble-modeller (kombinera flera algoritmer)
 - Prediktionsmarknad (köp/sälj prediktioner)
-

13. KÄLLOR OCH REFERENSER

13.1 Datakällor

ATG (2026). *Startlistor och resultat*. Tillgänglig: <https://www.atg.se> [Hämtad: 2026-02-07]

Svensk Travsport (2026). *Hästar, kuskar och tränare*. Tillgänglig: <https://www.travsport.se> [Hämtad: 2026-02-07]

Solvalla (2026). *Tävlingskalender och travprogram*. Tillgänglig: <https://www.solvalla.se> [Hämtad: 2026-02-07]

13.2 Metodologiska referenser

Kahneman, D. & Tversky, A. (1979). *Prospect Theory: An Analysis of Decision under Risk*. Econometrica, 47(2), pp. 263-291.

Silver, N. (2012). *The Signal and the Noise: Why So Many Predictions Fail – But Some Don't*. New York: Penguin Press.

James, G., Witten, D., Hastie, T. & Tibshirani, R. (2013). *An Introduction to Statistical Learning*. New York: Springer.

13.3 Domänkunskap

Gratistravtips (2026). *Dagens Dubbel tips och resultat*. Tillgänglig: <https://www.gratistravtips.se/dagens-dubbel> [Hämtad: 2026-02-07]

Travcash (2026). *Dagens Dubbel tips och spelform*. Tillgänglig: <https://travcash.se/dagens-dubbel/> [Hämtad: 2026-02-07]

14. BILAGOR

Bilaga A: Initiala viktningsparametrar (full tabell)

(Detaljerad viktning per faktor – inkluderas i implementationsdokument)

Bilaga B: Exempel på datainsamlingsformulär

(Mall för manuell datainsamling om automatisering ej möjlig)

Bilaga C: Statistiska metoder

(Fördjupning i korrelationsanalys, regressionsmodeller, etc.)

Bilaga D: Kodexempel

(Python-snippets för viktberäkning och prediktion)

15. GODKÄNNANDE OCH ÄNDRINGSHISTORIK

Version 1.0 – 7 februari 2026

Initial PRD skapad av Kent Lundgren

Nästa revidering: Efter första månaden (vecka 5, mars 2026)

Ändringslogg:

- [Datum]: [Beskrivning av ändring]
-

Dokumentslut

Detta PRD är ett levande dokument och kommer att uppdateras kontinuerligt baserat på systemets utveckling och resultat.