

DATA PROJECT CANVAS

Designed by: นวพร, พากษา, พัชราภา Date: 14/12/2568

Title: การวิเคราะห์ปัจจัยที่มีผลต่อการเกิดโรคเบาหวาน (Diabetes Risk Insight : A Data-Driven Screening Profiler)

1. Problem Statement/Background	2. Questions/Hypothesis	3. Value Proposition	4. Data Sources/Attributes
<ul style="list-style-type: none">ในประเทศไทยมีผู้ป่วยเป็นเบาหวานสูงกว่า 6.5 ล้านคน และเพิ่มขึ้นถึง 300,000 รายต่อปี โดยกว่า 40% ไม่รู้ตัวว่าแต่ตอนนี้เป็นโรคเบาหวาน จะเห็นได้ว่าในปัจจุบัน โรคไม่ติดต่อเรื้อรังเป็นปัจจุบันที่สำคัญของประเทศไทย	<ul style="list-style-type: none">ปัจจัยทางสุขภาพใดบ้างที่มีอิทธิพลต่อการเกิดโรคเบาหวานวิเคราะห์ลักษณะเฉพาะและสร้างแบบที่ตัดกรอง (Risk Profile) สำหรับผู้ที่มีความเสี่ยงปัจจุบันโรคเบาหวานในอนาคต	<ul style="list-style-type: none">สามารถประมวลผลความเสี่ยงได้ตามตัวเองเมื่อลงต้นตรวจสอบ Data Type, Missing Values, ค่า Null, ข้อมูลซ้ำ และนำไปสู่การปรับเปลี่ยนลบช่วงลักษณะลับๆ กรณีที่มีค่าเป็น 0 เป็น Missing Valuesคอลัมน์ Glucose, BloodPressure, SkinThickness, BMI ที่มีค่าเป็น 0 แทนค่าตัวอย่างค่าที่ไม่มีข้อมูลโดยแยกตามกลุ่ม Diabetes(0 หรือ 1)ให้สืบ dataset 401 row	<ul style="list-style-type: none">Data sources & collection<ul style="list-style-type: none">Kaggle : Healthcare Diabetes Dataset(2768 row)Data cleaning & preprocessing<ul style="list-style-type: none">- เปลี่ยนชื่อ colum ให้เป็น Diabetes- ตรวจสอบ Data Type, Missing Values, ค่า Null, ข้อมูลซ้ำ- ตรวจสอบ Data Type, Missing Values, ค่า Null, ข้อมูลซ้ำ- ลบช่วงลักษณะลับๆ กรณีที่มีค่าเป็น 0 เป็น Missing Values- คอลัมน์ Glucose, BloodPressure, SkinThickness, BMI ที่มีค่าเป็น 0 แทนค่าตัวอย่างค่าที่ไม่มีข้อมูลโดยแยกตามกลุ่ม Diabetes(0 หรือ 1)Target variables & feature<ul style="list-style-type: none">- Features : Glucose, Age, Skin Thickness, Blood Pressure, Insulin, BMI, Diabetes Pedigree Function- Target : Diabetes (1 = Diabetes, 0 = No Diabetes)Encoding & scaling strategies<ul style="list-style-type: none">- ทำ Log Transformation, Scaling ข้อมูล(Standard Scaler)
5. Analysis/Model Development	6. Findings and Insights	7. Recommendation/Action and Impact	
<ul style="list-style-type: none">ปัจจัยหลักที่มีผลต่อการเกิดโรคเบาหวาน ได้แก่ Glucose, Insulin, BMI, Age และร้อยร่วม ๆ กับ DiabetesPedigreeFunction, BloodPressure, SkinThicknessระดับน้ำตาลในเลือด(Blooducose) เป็นตัวชี้วัดที่สำคัญ ได้ที่สุดในการจำแนกการเป็นโรคเบาหวานอายุที่มากขึ้น(Age) และ ภาวะผู้ชายที่มีความเสี่ยงการเป็นโรคเบาหวานควรใช้ช่วงลักษณะดัชนีตาลในเลือด(Glucose)และระดับอัมูลิน(Insulin) ที่เจริญเติบโตในภูมิภาคเป็นโรคเบาหวาน	<ul style="list-style-type: none">Logistic Regression Model : Accuracy 76%, Recall (Class 1) 82%K-Nearest Neighbors Model : Accuracy 78%, Recall (Class 1) 58%Logistic Regression Model มีประสิทธิภาพต่ำกว่า K-Nearest Neighbors Model เพื่อระบุความสามารถในการจัดแบ่งระหว่างผู้ที่มีความเสี่ยงและผู้ที่ไม่มีความเสี่ยง	<ul style="list-style-type: none">ปรับแต่งการทำนายอย่างต่อเนื่อง หรือ ตรวจสอบค่าของโรคควรตรวจสอบค่าของ BMI ที่ต้องการลดลงควรออกกำลังกายอย่างประจำ เช่น กีฬา โยคะ หรือ ตรวจคัดกรองโรคเบาหวาน โดยแพทย์ผู้เชี่ยวชาญหากมีประวัติพ่อนในครอบครัวเป็นโรคเบาหวาน ควรดูแลตัวเองเป็นอย่างดี	