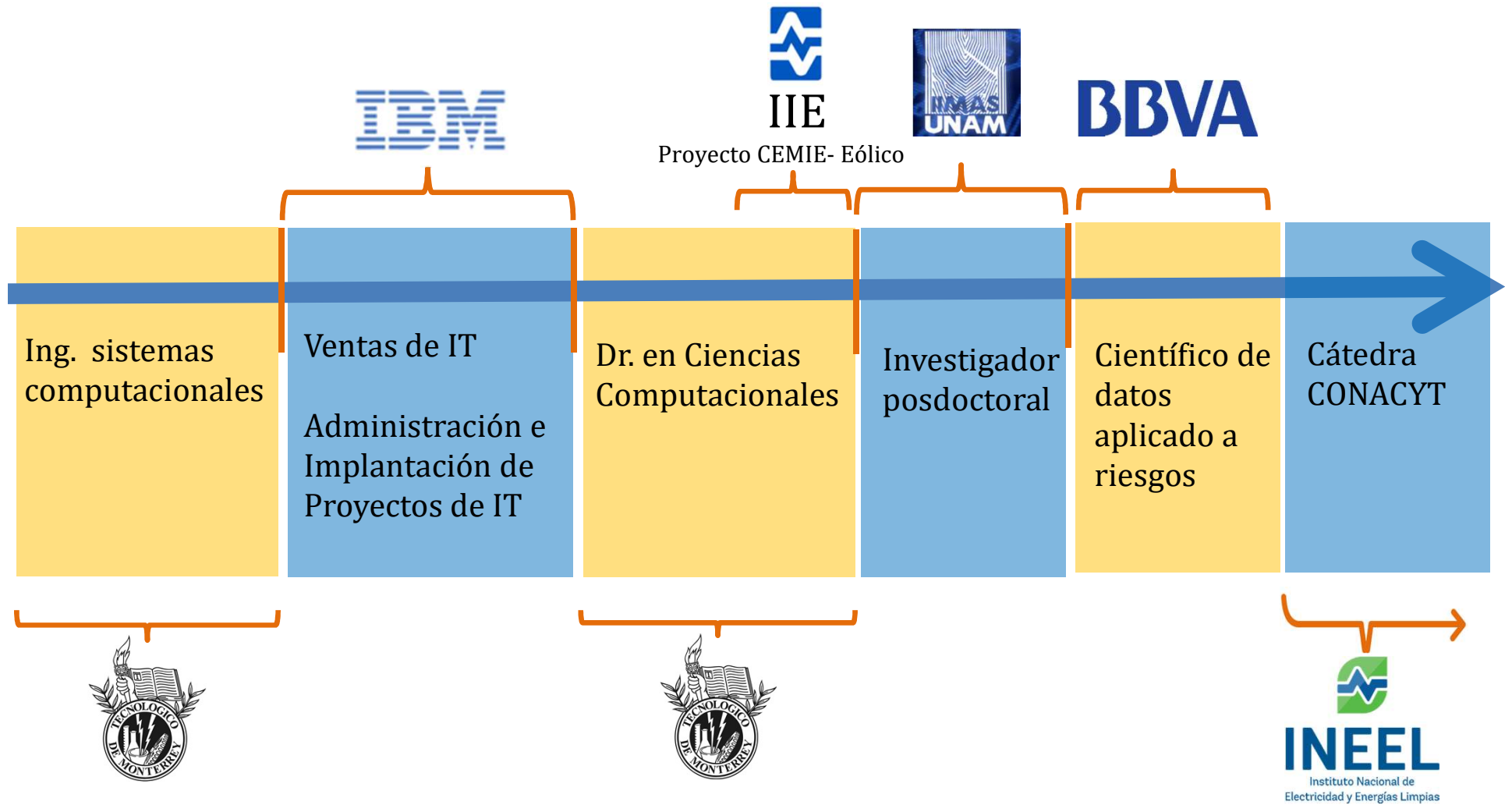


00: **Presentación**

Dr. Guillermo Santamaría-Bonfil

Formación



Temario

3 Semanas x 8 horas (3+5) = 24

1. Intro a la Ciencia de Datos
2. Intro al Pre-procesamiento
 1. Exploración Visual de Datos.
3. Regresión Lineal Simple y Múltiple.
 1. Pre-procesamiento: Normalización, Escalamiento, Transformación Logarítmica.
 2. Sel. Variables: Correlación y Mapas de Calor.
 3. Evaluación: Error Cuadrático Promedio (MSE), R^2 .
4. Clasificación: Regresión Logística, K-Vecinos Cercanos, Árboles de Decisión.
 1. Preprocesamiento: variables nominales con Peso de la Evidencia (WoE).
 2. Sel. Variables: Information Value.
 3. Evaluación: Matriz de Confusión y curvas ROC.
5. Pre-procesamiento: De texto a Bolsa de Palabras/ TF y TF-IDF.
 1. DataViz: Nube de Palabras.

Actividad-Evaluación

0. Ser autodidactas.

1. Regresión Múltiple para predicción de nuevos casos del COVID19 para el día siguiente

2. Clasificación Múltiples Clases.

3. Filtro SPAM aplicando diversos clasificadores y múltiples medidas de evaluación.