### Problem Statement

Hospitals struggle to continuously monitor patients' vital signs and promptly identify those at risk of adverse events. **Current systems often miss complex patterns or generate false alarms, leading to delayed interventions, increased risk, and higher expenses.**

### Proposed Solution

A mobile and desktop **application that notifies staff of patients needing assistance**. In an official capacity, this could vary by installation to include technologies such as pagers, intercom systems and audible alarms.

Behind the interface, a mix of **research and clinical data is joined with historical patient context**, creating a foundation where real-time data can be evaluated for anomalies.

| **Technical Solution** | **Purpose** | **Outcome** |
| --- | --- | --- |
| **Time Series Analysis**  ARIMA | Model vital signs to  predict future values | Establish baseline for health |
| **Supervised Learning**  Neural Networks, Random Forest, Gradient Boosting | Predict likelihood of  adverse events | Expand ARIMA output  with external factors (demographics, history) |
| **Unsupervised Learning**  UMAP, KMeans, DBSCAN | Identify unexpected deviations or patterns in patient data | Improve support for highly dimensional data |

### Data Sources

* **Publicly available:** MIMIC-III, MIMIC-IV, eICU, PhysioNet Databases
* **Generated, Synthetic:** As needed, mock values and context

Adequate **volume and dimensionality in data** is *critical* to legitimacy of this solution, and is expected to be **the biggest challenge for development.** Multiple datasets/databases are available for research, and will be the foundation of this capstone. After analysis, additional synthetic data can be used to simulate anomalies, detection, and client-side features (alerting).

Data extraction and transformation will include the **removal** of records which are considered **outliers (IQR)**, or have **significant value discrepancies** that may impact averaging (means) or normalization. The result of transformation should provide a diverse, high quality data set.

**Accessing this data** will be done through **RESTful API endpoints** as a first choice. Exports and uploads may be required for content that is not readily available by API. Similarly, should data requirements fall short, mock data from Kaggle (or similar) will be used to supplement any shortcomings.

### Differentiators & Goals

* **Scale & Scope:** Each integration is tailored to specific departments or facilities.
* **Comprehensive:** Combine leading medical research and clinical experience, with localized and external factors, as a baseline for every patient - 24/7/365.
* **Adaptive learning:** Discover new patterns and trends without downtime; bridging the gap and providing coverage between long-term support updates.
* **Compliance:** Centralized computing reduces distribution of confidential patient data.

The **primary goal** of this solution is to establish the previously mentioned models with a repeatable process, capable of reproducing and maintaining their efficiency.

The **secondary goal** is to apply real-time data streams, providing an indication for anomalies.

The **tertiary goal** is not strictly required, but would benefit demonstrations; an interface to facilitate various scenarios, with mock data and notification options.

#### *Applied Goals (Sample)*

* **Patients:** Faster awareness, quicker intervention
* **Patients:** Earlier intervention reduces risk, rate of complications
* **Staff:** Faster charting and leading indicators
* **Staff:** Larger knowledge base, consistent support

Together, these goals provide ample evidence in support of the solution solving a unique problem. Model selection and technical approach would be emphasized throughout, justifying the solution as a plausible enterprise application.