Financial context

- **Swaps**: Financial derivatives exchanging floating rates against fixed rate during \( T \) years (tenor).
  - The swap rate corresponding to the fixed rate \( S(T) \) such that the initial swap price = 0.
  - The forward swap rate \( F(U, T) \) is the swap rate \( S \) but for a swap starting at a future date \( U \).

- **Swaption**: Right to enter into a swap at time \( U \) (expiry) with predefined strike \( K \).
  - Can be seen as a call/put option on the underlying forward swap rate.
  - Implied volatility \( \sigma \) matching Black model price with market price.

- **Swaption (implied volatility) cube**: \( \sigma \) defined along three dimensions \((U, T, K)\).
  - Price convex w.r.t \( K \) and increasing w.r.t \( \sigma \) → strong no-arbitrage relationships.
  - Cash flows overlapping for close \((U, T)\) → statistical arbitrage.

Problem at hand

- **HSBC trading desks face data access constraints**: Swap smile - data providers refresh implied volatilities once a day.
  - Forward swap rates: refreshed on an intraday basis.
  - Traders would like to have an intraday view of implied volatilities.

- **Connection between several academic fields**: Financial mathematics: no-arbitrage relationship, dynamics for implied volatilities, Data science: features engineering (missing data, ...), model selection, hyperparameters calibration, Computer science: dealing with datasets size, calibration speed, ...

Formalization of the problem

- **Input data**: Past and current forward swap rate surfaces.
  - Past swaption implied volatility cubes.
- **Output data**: Current implied volatility cube.

With \( d_f \) the number of tenor points, \( d_i \) the number of expiry points, \( d_k \) the number of strike points, and \( n_d \) the number of days in our history:

Our learning problem amounts to finding an application

\[
 f : \mathbb{R}(d_f-1) \times d_f \times d_i \times d_i \times d_k \times d_k \times n_d \rightarrow \mathbb{R} \rightarrow \mathbb{R} \times d_f \times d_i \times d_k
\]

that minimizes the error criterion

\[
 \min_{(n_f,n_t)} \sum_{i=0}^{d_f} \sum_{j=0}^{d_i} \sum_{k=0}^{d_k} (Y^{h,i,j} - f(X^{h,i,j}))^2
\]

\[ f(X^{h,i,j}) : \text{Swaption implied volatility prediction with cube coordinate } h, i, j. \]

Simplified problem

- **Input data**: Today’s and yesterday’s forward swap rate surfaces.
  - Yesterday’s swaption implied volatility cube.

Modified error criterion:

\[
 \frac{1}{n_{\text{Scenarios}}} \sum_{\text{scenarios}} \sum_{h=0}^{d_f-1} \sum_{i=0}^{d_i} \sum_{j=0}^{d_k} (Y^{h,i,j} - f(X^{h,i,j}))^2
\]

\[ \text{(1)} \]

Numerical results

Heatmaps representing the root mean squared absolute error of swaption implied volatilities as a function of the log-moneyness \((\log \frac{U}{0.5})\) and \((U, T)\):

**Next steps**

- Increasing w.r.t \( T \), decreasing w.r.t \( U \).
- Outliers on the boundary (deep in the money).
- Skew dynamics approach more accurate.
- Nearest neighbors regressions approach smoother on the boundary.

### Error patterns common to the 3 schemes:

- Increasing w.r.t \( T \), decreasing w.r.t \( U \).
- Outliers on the boundary (deep in the money).
- Skew dynamics approach more accurate.
- Nearest neighbors regressions approach smoother on the boundary.

Absolute testing errors in captions are given by Eq. (1).

### Next steps

- Generating a viable implied volatility cube:
  - Statistical arbitrage over \((U, T)\) can be dealt with by regularization or validation procedures.
  - Stratifying estimation with market regimes (hidden markov models, ...).
  - Using autoencoder to capture an invariant structure of implied cube across time.

- Initial learning problem should be adapted to a time series setup:
  - Backtesting on real market data.
  - Dealing with missing/dirty data.
  - Recurrent Neural Network (LSTM).