Day 3 – Hybrid Machine Learning + Geostatistics

Today's Goal

Integrate:

- Machine learning
- Spatial geostatistics
- Hybrid prediction modelling

Load Data

```
library(readr)
hybrid_ml_geo <- read_csv("../data/hybrid_ml_geostatistical_data.csv")
head(hybrid_ml_geo)
```

```
# A tibble: 6 x 11
     id   utm_x    utm_y n_tested n_pos prevalence_true ml_signal_true
  <dbl>   <dbl>    <dbl>    <dbl> <dbl>           <dbl>          <dbl>
1     1 279090.  838870.       63     0         0.00167          -5.04
2     2  57751.  672968.      125     0         0.000452         -6.60
3     3 160759. 1365476.      139     0         0.00694          -3.80
4     4 949111. 1230481.      126     0         0.000334         -7.10
5     5 434005.  682490.       97     0         0.00205          -5.51
6     6 -91406.  925776.      122     3         0.00767          -3.16
# i 4 more variables: S_residual_true <dbl>, elevation <dbl>, ndvi <dbl>,
```