EXECUTIVE SUMMARY: Mental Health Workplace Analytics

Objective: The objective of this project was to develop machine learning models capable of predicting both the current presence of a mental health condition and the likelihood of seeking professional treatment. In addition, the analysis aimed to identify the most influential risk drivers and segment employees into distinct behavioral profiles to inform targeted and data-driven organizational interventions.

Dataset: 1,433 tech professionals, 69 engineered features, 26.3% missingness (structured imputation applied).

Prevalence: Among the surveyed employees, 52% reported a current mental health condition, and 58.5% had sought professional treatment. There is a strong correlation between having a condition and seeking treatment, though treatment is not universal, indicating potential barriers such as stigma or limited visibility of available support.

Model Performance:

Mental health condition prediction achieved strong performance, with Logistic Regression delivering the best results (ROC-AUC 0.92), followed by Random Forest (ROC-AUC 0.89). The most influential predictors were past disorder history, clinical diagnosis, family history, and reported productivity impact.

Treatment-seeking prediction showed similarly high accuracy, with Logistic Regression reaching a (ROC-AUC 0.92) and Random Forest (ROC-AUC 0.88). The strongest drivers of treatment behavior were clinical diagnosis, awareness of available support resources, workplace openness, and perceived stigma

Segmentation (KMeans k=3):

1. Clinically Diagnosed: high diagnosis, untreated interference.
2. Undiagnosed but Impacted: productivity loss without diagnosis.
3. Lower Disclosure: potential hidden risk.

Business Implications: Mental health outcomes are structurally influenced by diagnosis history and workplace support perception. A large undiagnosed segment shows measurable productivity impact.

Recommended Actions: Organizations should implement early screening to detect risk proactively, increase visibility of available mental health resources, normalize open conversations to reduce stigma, and apply targeted interventions tailored to distinct employee segments.

Conclusion: Mental health outcomes in tech workplaces are structurally influenced by diagnosis history, workplace openness, and support visibility.

The predictive signal is strong and segmentation is actionable, enabling data-driven interventions to reduce untreated cases and productivity loss.