

N-Gram-Based Text Categorization

William B. Cavnar and John M. Trenkle
Environmental Research Institute of Michigan
P.O. Box 134001
Ann Arbor MI 48113-4001