

transformer variants at init by hidden dim

